THE METHOD OF PATH COEFFICIENTS

By

SEWALL WRIGHT
Department of Zoology, The University of Chicago.

Introduction

The method of path coefficients was suggested a number of years ago (Wright 1918, more fully 1920, 1921), as a flexible means of relating the correlation coefficients between variables in a multiple system to the functional relations among them. The method has been applied in quite a variety of cases. It seems desirable now to make a restatement of the theory and to review the types of application, especially as there has been a certain amount of misunderstanding both of purpose and of procedure.

Basic Formulae

The object of investigation is a system of variable quantities, arranged in a typically branching sequential order representative of some chosen point of view toward the functional relations. Such a system is conveniently represented in a diagram such as Fig. 1. Those variables which are treated as dependent are connected with those of which they are considered functions by arrows. The system of factors back of each variable may be made formally complete by the introduction of symbols representative of total residual determination (as V_u in Fig. 1). A residual correlation between variables is represented by a double-headed arrow. It will be assumed that all relations are linear.[1] Thus each variable is related to those from which uni-

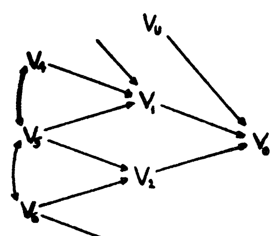


FIG. 1

[1] Relations which are far from linear with respect to the absolute values of the variables may be approximately linear with respect to variations, if the coefficients of variability are small. Thus if $V_0 = f(V_1, V_2, \cdots, V_n)$,

directional arrows are drawn to it by an equation of the following type, where $(V_0 - \bar{V}_0)$, $(V_1 - \bar{V}_1)$, ... etc. represent deviations from the means and c_{01}, c_{02} etc. are the coefficients.

$$(1) \qquad (V_0 - \bar{V}_0) = c_{01}(V_1 - \bar{V}_1) + c_{02}(V_2 - \bar{V}_2) + \cdots + c_{0n}(V_n - \bar{V}_n).$$

It is convenient to measure the deviation of each variable by its standard deviation. Let $X_0 = \frac{V_0 - \bar{V}_0}{\sigma_0}$, $X_1 = \frac{V_1 - \bar{V}_1}{\sigma_1}$ etc., and let $p_{0i} = c_{0i} \frac{\sigma_i}{\sigma_0}$,

$$(2) \qquad X_0 = p_{01} X_1 + p_{02} X_2 + \cdots + p_{0n} X_n .$$

The coefficients in this form are of the type called path coefficients. Each obviously measures the fraction of the standard deviation of the dependent variable (with the appropriate sign) for which the designated factor is directly responsible, in the sense of the fraction which would be found if this factor varies to the same extent as in the observed data while all others (including residual factors X_u) are constant. This definition (except for determination of sign) can be written as follows, putting the constant factors after a dot.

$$(3) \qquad p_{01} = \frac{\sigma_{0 \cdot 23 \cdots n, u}}{\sigma_0} \cdot \frac{\sigma_1}{\sigma_{1 \cdot 23 \cdots n, u}} .$$

It is sometimes convenient to represent the standard deviation due directly to a particular factor by a symbol. The form $\sigma_{0(1)} = p_{01} \sigma_0$ will be used. Obviously $\sigma_{0(1)} = c_{01} \sigma_1$ and, neglecting sign,

$$(4) \qquad c_{01} = \frac{\sigma_{0 \cdot 23 \cdots n, u}}{\sigma_{1 \cdot 23 \cdots n, u}}$$

The theorem which makes the path coefficient useful in relating correlations to functional relation is a very simple one. The correlation between V_0 and any other variable V_q in such a system as Fig. 2 can be written in the form

---

the relation of small deviations from the mean values are approximately the first order terms of an expansion by Taylor's Theorem. The error may be represented by a residual term R.

$$\delta V_0 = \frac{\partial V_0}{\partial V_1} \delta V_1 + \frac{\partial V_0}{\partial V_2} \delta V_2 + \cdots + R.$$

$$r_{0q} = \frac{1}{N}\Sigma X_0 \cdot X_q = \frac{1}{N}\Sigma X_q (P_{01} X_1 + P_{02} X_2 + \cdots + P_{0n} X_n) \tag{5}$$
$$= P_{01} r_{q1} + P_{02} r_{q2} + \cdots + P_{0n} r_{qn}$$
$$= \Sigma P_{0i} r_{qi} .$$

The correlation is thus analyzed into contributions from all of the paths in the diagram (Fig. 2) passing through each factor of one of the variables.

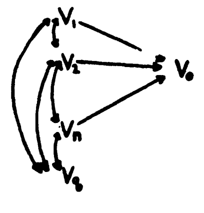


FIG. 2

But the correlation terms symbolized by r_{qi} may be capable of analysis by application of this same formula. By repeated analysis of this sort, as far as the diagram (such as Fig. 1) permits, we are led to the following principle: Any correlation between variables in a network of sequential relations can be analyzed into contributions from all of the paths (direct or through common factors) by which the two variables are connected, such that the value of each contribution is the product of the coefficients pertaining to the elementary paths. If residual correlations are present (represented by bidirectional arrows) one (but never more than one) of the coefficients thus multiplied together to give the contribution of a connecting path, may be a correlation coefficient. The others are all path coefficients

In tracing connecting paths it is obvious that one may trace back along the arrows and then forward as well as directly from one variable to the other (perhaps through intervening variables) but never forward and then back. That two factors affect the same dependent variable does not contribute to the correlation between them. Similarly two variables which are correlated with a third are not necessarily correlated with each other. As illustrations of these principles consider the correlations between some of the variables in Fig. 1.

$$r_{36} = P_{36} \qquad r_{46} = 0 \qquad r_{13} = P_{15} \cdot r_{56} \cdot P_{36}$$
$$r_{23} = P_{25} \cdot r_{56} \cdot P_{36} + P_{26} \cdot P_{36} \qquad r_{12} = P_{14} \cdot r_{45} \cdot P_{25} + P_{15} \cdot P_{25} + P_{15} \cdot r_{56} \cdot P_{26}$$

It is sometimes convenient to use an extension of the symbolism in dealing with compound paths.

$$r_{02} = p_{02} + p_{01\dot{5}2} + p_{01\overline{45}2} + p_{01\overline{56}2}$$

$$r_{05} = p_{015} + p_{025} + p_{01\overline{45}} + p_{02\overline{65}}$$

In this symbolism all of the variables along a contributing path are listed in proper order. If the path passes through a represented common factor, the latter is indicated by a dot. If it involves an unanalyzed correlation the two ultimate correlated variables may be indicated by a line as above. The evaluation of such compound path coefficients is obvious,

$p_{015} = p_{01}\, p_{15}$, $\qquad p_{01\dot{5}2} = p_{01}\, p_{15}\, p_{25}$

$p_{01\overline{45}2} = p_{01}\, p_{14}\, r_{45}\, p_{25}$, $\qquad p_{01\overline{45}} = p_{01}\, p_{14}\, r_{45}$, etc.

It is to be noted that the symbolism does not apply to the indicated variables in an absolute sense but is always to be understood as relative to a particular arrangement of the variables, i.e. to a particular point of view with respect to the functional relations.

A special case of equation (5) arises if one correlates a variable with itself, taking into account *all* factors (known and unknown)

(6) $$r_{00} = \Sigma\, p_{0i}\, r_{0i} = 1.$$

This may be put in a form which is usually more convenient by further analysis of $r_{0i} = p_{0i} + \Sigma\, p_{0j}\, r_{ij}$

(7) $$\Sigma\, p_{0i}^2 + 2\,\Sigma\, p_{0i}\, p_{0j}\, r_{ij} = 1.$$

Degree of Determination

From the formula $p_{01}^2 = \dfrac{\sigma_{0(1)}^2}{\sigma_0^2}$ it is obvious that a squared path coefficient measures the portion of the variance of the dependent variable for which the independent variable is directly responsible, under the point of view adopted. The squared path coefficient may accordingly be called a coefficient of determination. Such coefficients were used before the term path coefficient was applied to the square root. (Wright 1918.)

The sum of the squared path coefficients is unity only in the case in which there are no correlations among the factors. It is necessary, therefore, to recognize additional terms measuring the changes in variance (positive or negative) due to correlated occurrence of the contributions of such factors ($2 P_{oi} P_{oj} r_{ij}$, etc. in equation 7). It is tempting to apportion determination among the factors by using the terms $P_{oi} r_{oi}$ of equation (6) as measures of determination, and this has been done by some authors, e.g. Kirchevsky (1927) who independently reached a somewhat similar viewpoint on the interpretation of systems of correlated variables in other respects. No transparent meaning can be attached to such expressions (which may be negative). The term does not measure direct determination since it involves indirect connections between the variables. Neither does it measure total determination, direct and indirect. This is given by the squared correlation coefficient.

The Correlation between Linear Functions

The most direct application of the method is in the estimation of the correlation between two variables which are functions (in part at least) of the same variables. Let V_S and V_T be two variables whose correlation is desired.

$$V_S = c_S + c_{S1} V_1 + c_{S2} V_2 + \cdots + c_{Si} V_i$$

$$V_T = c_T + c_{T1} V_1 + c_{T2} V_2 + \cdots + c_{Ti} V_i$$

$$(8) \quad \sigma_S^2 = \Sigma c_{Si}^2 \sigma_i^2 + 2 \Sigma c_{Si} c_{Sj} \sigma_i \sigma_j r_{ij}$$

$$\sigma_T^2 = \Sigma c_{Ti}^2 \sigma_i^2 + 2 \Sigma c_{Ti} c_{Tj} \sigma_i \sigma_j r_{ij}$$

$$P_{Si} = c_{Si} \frac{\sigma_i}{\sigma_S}, \qquad P_{Ti} = c_{Ti} \frac{\sigma_i}{\sigma_T}$$

$$(9) \quad r_{ST} = \Sigma P_{Si} P_{Ti} + \Sigma P_{Si} r_{ij} P_{Tj}$$

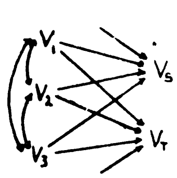


FIG. 3

As an example, suppose that we wish to find the correlation between the compound variables V_S and V_T where $V_S = V_1 + V_2 + V_3$ and $V_T = V_1 + V_2 + V_4$ knowing that V_1, V_2, V_3 and V_4 are all of equal variability ($\sigma_1 = \sigma_2 = \sigma_3 = \sigma_4$)

and are independent $(r_{12} = r_{13} = r_{14} = r_{23} = r_{24} = r_{34} = 0)$

$$\sigma_S^2 = \sigma_T^2 = 3\,\sigma_1^2$$

$$p_{ST} = p_{S2} = p_{S3} = p_{T1} = p_{T2} = p_{T4} = \sqrt{1/3}$$

$$r_{ST} = p_{S1}\,p_{T1} + p_{S2}\,p_{T2} = 2/3 .$$

Again suppose that we wish to estimate the true correlation between two variables from that between measurements known to be subject to considerable random error. Assume that the correlation between two measurements of the same variate has been found in each case. It is instructive to work this out from two different points of view.

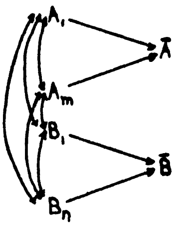


FIG. 4

Let $\bar{A}$ be the mean of m 'measurements $(A_1, A_2, \cdots, A_m)$ of A. Let $\bar{B}$ be the mean of n measurements $(B_1, B_2, \cdots, B_n)$ of B.

The known correlations are those between measures of A (r_{AA}), between measures of B (r_{BB}) and between measures of A and B (r_{AB}). Expressing the complete determination of $\bar{A}$ and $\bar{B}$ by their components, using equation (6): $\Sigma\, p_{\bar{A}A}\, r_{\bar{A}A} = m\, p_{\bar{A}A}^2 \left[1 + (m-1)\, r_{AA}\right] = 1$

$$\Sigma\, p_{\bar{B}B}\, r_{\bar{B}B} = n\, p_{\bar{B}B}^2 \left[1 + (n-1)\, r_{BB}\right] = 1$$

From these $p_{\bar{A}A} = \sqrt{\dfrac{1}{m[1+(m-1)r_{AA}]}}$, $p_{\bar{B}B} = \sqrt{\dfrac{1}{n[1+(n-1)r_{BB}]}}$

The correlation between $\bar{A}$ and $\bar{B}$ can be written

$$r_{\bar{A}\bar{B}} = m\, p_{\bar{A}A}\, r_{A\bar{B}} = m\cdot n\cdot p_{\bar{A}A}\cdot p_{\bar{B}B}\cdot r_{AB}$$

(10)
$$= r_{AB}\sqrt{\frac{m\cdot n}{[1+(m-1)r_{AA}][1+(n-1)r_{BB}]}} .$$

For indefinitely large values of m and n the averages may be considered as true scores, A_T and B_T.

(11)
$$r_{A_T B_T} = \frac{r_{AB}}{\sqrt{r_{AA}\cdot r_{BB}}} .$$

This result can be reached much more directly by the simpler set up (Fig. 5) in which the observed measurements are represented as functions of the true scores A_T, B_T and of random

errors. Note that the directions of the arrows are the reverse of those in Fig. 4.

$$r_{AA} = p^2_{AA_T} \quad \text{giving} \quad p_{AA_T} = \sqrt{r_{AA}}$$

$$r_{BB} = p^2_{BB_T} \quad \text{giving} \quad p_{BB_T} = \sqrt{r_{BB}}$$

$$r_{AB} = p_{AA_T}\, r_{A_TB_T}\, p_{BB_T}$$

again giving $$r_{A_TB_T} = \frac{r_{AB}}{\sqrt{r_{AA}\cdot r_{BB}}}.$$

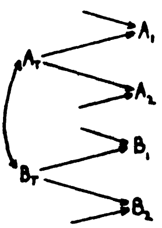


FIG. 5

This formula is, of course, Spearman's correction for attenuation. The purpose here is to bring out the simple way in which such formulae can be obtained by the method of path coefficients. The following is a more complex case in which a simple method is more essential.

The Statistical Effects of Inbreeding[2]

Assume for simplicity that the effects of different genetic factors combine additively (no dominance or epistasis). In Fig. 6, P_1 and P_2 represent the genetic constitution of two parents and O of their offspring. The constitution of the latter (under the above assumption and ignoring the possibility of sex linkage) is equally and completely determined by the constitutions of the two germ cells (G_1, G_2) which united to produce it. It will be convenient to represent the path coefficients and correlations by single letters:

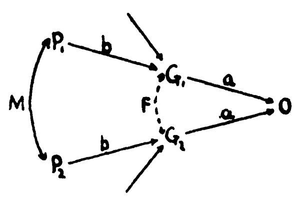


FIG. 6

$$a = p_{OG_1} = p_{OG_2}, \qquad b = p_{G_1P_1} = p_{G_2P_2},$$

$$F = r_{G_1G_2}, \qquad M = r_{P_1P_2}$$

The determination of O by G_1 and G_2 can be expressed in

[2] The purpose in presenting this and later examples is to illustrate something of the range of applicability of the method, rather than to give a detailed analysis of each case. For the latter the reader must be referred to the references cited at the end.

the equation $2a^2(1+F)=1$ (by equation 7) giving

(12) $$a=\sqrt{\frac{1}{2(1+F)}}$$

Two complementary germ cells (G_A, G_B) Fig. 7, such as could arise from the same reduction division have the same relation to the genetic constitution of the parent (which they completely determine in a mathematical sense) as the two germ cells which united to produce the parent, assuming no selection and that different series of allelomorphs are combined at random, (an assumption compatible with linkage among the genes and with inbreeding, but not with assortative mating). Using primes for the path coefficients and correlations of the preceding generations.

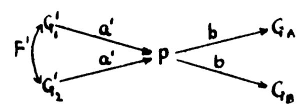


FIG. 7

(13) $$b = r_{PG'} = a'(1+F') = \sqrt{\frac{1+F'}{2}}$$

Since $a'=\sqrt{\frac{1}{2(1+F')}}$ (by equation 12, applied to the preceding generation) $ba'=1/2$ irrespective of correlation between the parents under the assumed conditions.

(14) $$F=b^2M, \qquad M=\frac{2F}{1+F'}$$

The correlation between uniting gametes is directly related to the percentage of heterozygosis. Below is the correlation table between uniting gametes in a population in which genes A and a are present in the frequencies of q and $1-q$ respectively and the proportion of heterozygotes is p. By the usual formula for correlation

(15) $$F=\frac{(q-p/2)-q^2}{q(1-q)} = 1-\frac{p}{2q(1-q)}$$

$$p=2q(1-q)(1-F).$$

All of the path coefficients and correlations have now been expressed in terms of F's. Various applications can be made. As a simple case consider the

	a	A	Total
A	$\frac{p}{2}$	$q-p/2$	q
a	$1-q-\frac{p}{2}$	$p/2$	$1-q$
Total	$1-q$	q	1

effects of continued brother-sister mating: Analyzing the correlation (M) Fig. 8, between the parents by tracing the connecting paths:

(16) $M = 2a'^2 b'^2 (1+M')$.

Expressing all coefficients in terms of $F's$ and reducing

(17) $F = 1/4\,(1+2F'+F'')$

(18) $P = \frac{p'}{2} + \frac{p''}{4}$.

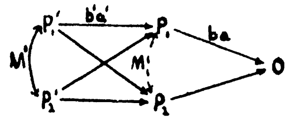


FIG. 8

Thus the percentage of heterozygosis (to which the effects of inbreeding are directly related) is very simply related to the percentages in the two preceding generations. If there were initially 50% i.e. ½ heterozygosis, that of later generations would be given by the terms of a series of fractions in which each numerator is the sum of the two preceding numerators (Fibonacci series) if the denominator is doubled in each generation. This rule was derived empirically by Jennings (1916) on working out in detail the consequences of every possible mating, generation after generation. The analysis by path coefficients (Wright 1921b) not only demonstrates the generality of the empirical rule but can be applied as easily to more complicated cases in which the analysis by types of mating would be practically impossible. Consider, for example, the more general case of a population restricted to N_m mature males and N_f mature females (Wright 1931b). Under random mating, the chance of a mating of full brother and sister is $\frac{1}{N_m \cdot N_f}$, of half brother and sister $\frac{N_m+N_f-2}{N_m \cdot N_f}$ and of less closely related individuals $\frac{(N_m-1)(N_f-1)}{N_m\, N_f}$.

The correlation between mating individuals is thus

(19) $M = a'^2 b'^2 \left[\frac{2+2M'}{N_m \cdot N_f} + \left(\frac{N_m+N_f-2}{N_m \cdot N_f}\right)(1+3M') + \frac{(N_m-1)(N_f-1)}{N_m \cdot N_f}\cdot 4M'\right]$

which yields on reduction

(20) $F = F' + \left(\frac{N_m+N_f}{8N_m \cdot N_f}\right)(1-2F'+F'')$

(21) $p = p'\left(\frac{N_m+N_f}{8N_m \cdot N_f}\right)(2p'-p'')$.

Equating $\frac{P}{P'}$ to $\frac{P'}{P''}$ gives $\left(\frac{1}{8N_m}+\frac{1}{8N_f}\right)$ as the approximate rate of reduction of heterozygosis per generation. The special case in which the population is equally divided between males and females $(N_m = N_f = \frac{N}{2})$ gives $\frac{1}{2N}$ as the rate of reduction, a figure recently verified by R. A. Fisher by a very different mode of analysis.

The method has also been applied in the much more complicated case of assortative mating based on somatic resemblance (Wright 1921b).

In the case of the irregular inbreeding encountered in live stock pedigrees (Wright 1922, 1923a), the basic formula of path coefficients leads immediately to the formula

$$F = \Sigma\left[\left(\tfrac{1}{2}\right)^{N_s+N_d+1}\cdot(1+F_A)\right], \tag{22}$$

where N_s and N_d are the number of generations from sire and dam respectively to the common ancestor (A) at the head of each connecting path. By appropriate sampling methods (Wright & McPhee, 1925) this formula can be used in the study of whole breeds. Closely allied is the formula for the genetic correlation between any two individuals (X, Y). Letting N and N' be the generations from X and Y respectively to the common ancestor of any connecting path

$$R_{XY} = \frac{\Sigma\left[\left(\tfrac{1}{2}\right)^{N+N'}\cdot(1+F_A)\right]}{\sqrt{(1+F_X)(1+F_Y)}}. \tag{23}$$

These formulae have been extensively applied in breed analysis (Wright 1923b-c, McPhee & Wright 1925, 1926, Smith 1926, Calder 1927, Lush 1932).

Multiple Regression

The preceding applications have consisted in the main in the deduction of correlation coefficients from knowledge of the functional relations. The method can be applied as well to the inverse problem, that of finding the best linear expression for one variable in terms of a number of others, from knowledge of the cor-

relation coefficients. No assumptions are made with respect to causal relations. Analysis of the correlations between V_0 and the other variables (Fig. 9), by the basic formula, gives the following set of equations.

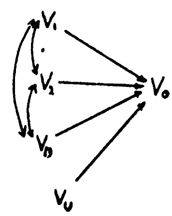


FIG. 9

$$
\begin{aligned}
r_{01} &= p_{01} + p_{02}\, r_{12} + \cdots + p_{0n}\, r_{1n} \\
r_{02} &= p_{01}\, r_{12} + p_{02} + \cdots + p_{0n}\, r_{2n} \\
&\text{- - - - - - - - - - - - - - - - -} \\
r_{0n} &= p_{01}\, r_{1n} + p_{02}\, r_{2n} + \cdots + p_{0n}
\end{aligned} \tag{24}
$$

Obviously these are merely the normal equations of the method of least squares in a slightly disguised form, as might be expected from the derivation of the basic formula. The solution for the path coefficients, expressed in terms of determinants, merely need to be multiplied by the proper ratio of standard deviations to give Pearson's formulae for the partial regression coefficients. The method of path coefficients here merely furnishes a convenient mnemonic rule for writing the normal equations.

The correlation between the actual values of V_0 and the estimates (V_0') (Fig. 10) from each set of values of the other variables, (given by the regression equation) is Pearson's coefficient of multiple correlation. Let V_u stand for the array of residual factors of V_0 in a form independent of the known factors. We may write an equation of complete determination (Fig. 9)

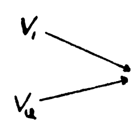


FIG. 10

$$\sum_1^n p_{0i}\, r_{0i} + p_{0u}\, r_{0u} = 1$$

$$\sum_1^n p_{0i}\, r_{0i} = 1 - r_{0u}^2 \quad \text{since } p_{0u} = r_{0u}.$$

But $\qquad r_{00'}^2 = 1 - r_{0u}^2$ (Fig. 10) $\qquad$ Therefore

$$r_{00'} = r_{0(12\cdots n)} = \sqrt{\Sigma\, p_{0i}\, r_{0i}}\,. \tag{25}$$

It is unnecessary to give illustrations of the use of the method in obtaining ordinary estimation or prediction equations.

A somewhat different type of application has been made in estimating the transmitting capacity of dairy sires (Wright 1932a). In this case the necessary correlations were deduced from Mendelian theory checked by observed correlations between the sire's female relatives and his daughters. These correlations were then used to calculate the multiple regression of sire on daughters and their dams.

Partial Correlation

It is sometimes of interest to find the values which statistics would take, on the average, in data selected for constancy of one or more variables.

$$(26) \qquad \sigma^2_{0\cdot 12\cdots n} = \sigma'^2_{0(u)} = \sigma^2_0 \cdot p^2_{0u} = \sigma^2_0 (1 - r^2_{0(12\cdots n)}),$$

a well known formula. Inspection of equations (3) and (4) and of the definition of $\sigma_{0(1)}$ gives the following for the standard deviation of V_0 due directly to V_1, under constancy of V_2 etc., and for the related path coefficient and concrete partial regression coefficient under the same conditions.

$$(27) \qquad \sigma_{0(1)\cdot 2\cdots m} = \sigma_{0(1)} \frac{\sigma_{1\cdot 2\cdots m}}{\sigma_1} = \sigma_{0(1)} \sqrt{1 - r^2_{1(2\cdots m)}}$$

$$(28) \qquad p_{01\cdot 2\cdots m} = \frac{\sigma_{0(1)\cdot 2\cdots m}}{\sigma_{0\cdot 2\cdots m}} = p_{01} \sqrt{\frac{1 - r^2_{1(2\cdots m)}}{1 - r^2_{0(2\cdots m)}}}$$

$$(29) \qquad c_{01\cdot 2\cdots m} = \frac{\sigma_{0(1)\cdot 2\cdots m}}{\sigma_{1\cdot 2\cdots m}} = c_{01}$$

As might be expected, the concrete coefficients of the multiple regression equation (c_{01} etc.) remain the same (on the average) in samples selected for constancy of one or more of the factors, while the abstract path coefficients are altered in value in such samples.

The formula for partial correlation can be derived from the formula $\sigma^2_{0\cdot 1} = \sigma^2_0 (1 - r^2_{01})$

as applied to the data in which particular variables $(V_2 \cdots V_m)$ are constant.

$$(30)\qquad \sigma^2_{0.12\cdots m} = \sigma^2_{0.2\cdots m}\,(1 - r^2_{01.2\cdots m})$$

$$r^2_{01.2\cdots m} = 1 - \frac{\sigma^2_{0.12\cdots m}}{\sigma^2_{0.2\cdots m}} = 1 - \frac{1 - r^2_{0(12\cdots m)}}{1 - r^2_{0(2\cdots m)}}\,.$$

This derivation leaves the sign uncertain but this is easily determined from a different approach. In Fig. 11, V_u includes *all* factors of V_0 other than the designated independent variable V_1 and the variables $V_2 \cdots V_m$ which are to be held constant. V_v represents the residual factor for V_1, in relation to the factors $V_2 \cdots V_m$. The value of P'_{01} in this system is not of course the same as P_{01} in the preceding discussion in which other variables than $V_2 \cdots V_m$ (those to be made constant) were treated as factors of V_0.

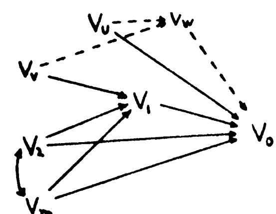


FIG. 11

Since
$$c_{01} = r_{01}\frac{\sigma_0}{\sigma_1}\,,\quad c_{01.2\cdots m} = r_{01.2\cdots m}\frac{\sigma_{0.2\cdots m}}{\sigma_{1.2\cdots m}}\,.$$

But $\quad c_{01.2\cdots m} = c_{01} = P_{01}\dfrac{\sigma_0}{\sigma_1}$ $\qquad$ Therefore

$$(31)\qquad r_{01.2\cdots m} = P_{01}\cdot\frac{\sigma_0}{\sigma_1}\cdot\frac{\sigma_{1.2\cdots m}}{\sigma_{0.2\cdots m}} = P_{01}\sqrt{\frac{1 - r^2_{1(2\cdots m)}}{1 - r^2_{0(2\cdots m)}}}\,.$$

Thus $\quad r_{01.2\cdots m}$ has the same sign as P_{01}.

This is simply formula 28 except that it is in a set up in which all factors of V_0 except V_1 are held constant, in which case $P_{01.2\cdots m}$ becomes $r_{01.2\cdots m}$.

Since
$$1 - r^2_{0(12\cdots m)} = P^2_{0u}$$
$$1 - r^2_{1(2\cdots m)} = P^2_{1v}$$
$$1 - r^2_{0(2\cdots m)} = P^2_{0u} + P^2_{01}P^2_{1v} \text{ OR } P^2_{0w}\,,$$

letting V_w represent the combination of V_u and V_v the above formulae for partial correlation can be written in a number of very compact forms.

$$(32)\qquad r_{01.2\cdots m} = \sqrt{1 - \frac{p^2_{ou}}{p^2_{ow}}}$$

$$(33)\qquad = \frac{p_{01} \cdot p_{1v}}{\sqrt{p^2_{ou} + p^2_{01} \cdot p^2_{1v}}}$$

$$(34)\qquad = \frac{p_{01} \cdot p_{1v}}{p_{ow}}\,.$$

The first of these is identical with 30.

Symbolism

The most widely current symbol for a partial regression coefficient is Yule's expression $b_{01.2\cdots m}$. Kelley (1923) uses a similar expression $\beta_{01.2\cdots m}$ for the coefficients in abstract form. These have an advantage over the symbols used here (c_{01}, p_{01} respectively) in that they define certain absolute functions of the variables, while the latter symbols have meaning only in relation to a particular arrangement. This relativity of meaning can not, however, cause confusion as long as one is dealing with only a single system. If the problem is of a more complex sort than the calculation of a prediction formula, the β symbolism becomes too cumbersome for convenience. The current symbolism has the further disadvantage of a certain lack of logical consistency. In the expression $\sigma_{0.12}$, $r_{01.23}$ and $b_{01.23}$ the subscripts to the right of the dot are understood to represent factors held constant. In the expressions $r_{0.12}$ and $\beta_{01.23}$ this is not the case. If we wish to represent the multiple correlation of X_0 with X_1 and X_2, independent of X_3, or the beta (path coefficient) for the influence of X_1 on X_0 in data involving also X_2 and X_3 but in which X_3 is held constant, it would apparently be necessary under the usual symbolism to write such ambiguous expressions $r_{0.123.3}$ and $\beta_{01.23.3}$ respectively. Pearson's method

of writing constant factors as subscripts to the left of the main symbol avoids these difficulties and is the one which I have followed in earlier papers. The dot symbolism has, however, become so firmly established in the cases of the standard deviation and correlation coefficient that it is probably best to recognize it as the general device for indicating constant factors and to replace it in those symbols in which it is used for a different purpose.

There is no difficulty in the case of multiple correlation. The expression $r_{0(12)}$ may be used for the correlation of X_0 with X_1 and X_2 jointly and the expression $r_{0(12)\cdot 3}$ is an unambiguous symbol for the multiple correlation independent of X_3.

As noted above, it is not desirable in the usual application of path coefficients to encumber the symbols with a list of the factors of which each dependent variable is treated as a function. This can be left to a diagram. Where a complete formal symbolism is desirable, the list of factors might follow a semicolon instead of a dot. Thus $P_{01:23.3}$ would unambiguously represent the path coefficient relating X_0 to X_1 in a system in which X_0 is treated as a function of X_1, X_2 and X_3 but in which X_3 is to be held constant. There is, however, little need for such complicated expressions.

Quantitative Evaluation of Causal Relations

While the method of path coefficients is directly applicable to such problems as the estimation of correlation coefficients from knowledge of the mathematical relations between variables, or the converse (multiple regression) it was developed primarily as a means of combining the quantitative information given by a system of correlation coefficients with such information as may be at hand with regard to the causal relations, and thus of making quantitative an interpretation which would otherwise be merely qualitative.

How far such causal analysis has meaning is a question on which there is difference of opinion. Some authors (Pearson, Niles) have contended that the designation of the relation be-

tween two variables as one of cause and effect involves a false conception; that we can merely observe more or less perfect correlation. This view seems to imply that direction in time is of no significance, and indeed G. N. Lewis has recently argued for the complete symmetry of the physicist's time. The common sense view that direction in time is a basic perception is not without support, however.

Under the theory of relativity, the elementary physical reality seems to be the point event located at a particular position in the space and time of a particular viewpoint. The objective world is to be thought of as a complex network of point events. Although two such events sufficiently remote from each other in space, relative to their separation in time, may have their order of succession in time reversed in the systems of two dfferent observers, order in time is invariant along any strand of this network involving continuity of physical action. Thus the succession of collisions suffered by a particular body or by a beam of light is the same to all observers. Such successions of events as involved in the movement of a shadow over a surface may indeed be reversed by change of viewpoint, if the shadow happens to be moving more rapidly than the velocity of light, but the continuity of physical action here is not along the path of the shadow but traces separately to each point in this path from the points of interception of the light. There is frequently difficulty in complex cases in distinguishing lines of direct causation from correlations due to common causation but in principle the distinction is clear enough. Experimental intervention is possible only in the true lines of causation.

In the world of large scale events, certain patterns tend to recur. Certain recurrent successions of events come to be recognized, experimentally or otherwise, as lines of causation in the above sense. Different lines of this character may come together in a certain type of event or may diverge from one. In many cases a fairly adequate representation of the course of nature can

be obtained by viewing it as a coarse network in which the "events" of interest are the deviations in the values of certain measurable quantities. A qualitative scheme depends on observation of sequences and experimental intervention. It is of interest to make such a scheme at least roughly quantitative in the sense of evaluating the relative importance of action along different paths. This was the primary purpose of the method of path coefficients.

Birth Weight of Guinea Pigs

The simplest application of this sort has been in connection with the factors which determine the weight of guinea pigs at birth (Wright 1921a). Minot (1891) noted that the average birth weight is smaller, the greater the size of the litter. He reasoned that this might be due either to a competition between the developing foetuses, or merely to an effect of a large litter in stimulating somewhat premature birth. In confirmation of the latter hypothesis he found that the gestation period was several days shorter in large litters than in small ones and that there was in fact a direct relation between length of gestation period and birth weight. After some discussion, he concluded that the data afforded no evidence of growth competition and thus he decided in favor of the second hypothesis. I was able to confirm Minot's observations, obtaining the following data in a large stock of guinea pigs. The mean birth weight (in grams) of the animals in the litter is the birth weight used. The interval between litters, where less than 75 days is approximately the gestation period. Standard errors are given.

	Mean	*S D*	
B (Birth weight)	82.24 ± 0.51	18.60 ± 0.36	$r_{BI} = +.533 \pm .020$
I (Interval)	68.93 ± 0.05	1.91 ± 0.04	$r_{BL} = -.658 \pm .010$
L (Size of litter)	2.91 ± 0.04	1.29 ± 0.03	$r_{IL} = -.457 \pm .022$

The correlation between birth weight and size of litter was based on 3353 cases, the other two correlations on 1317 cases.

In order to make a comparison of Minot's two alternatives,

these may be represented graphically in a single diagram.

Birth weight B is completely determined (in the mathematical sense rather than causally) by the prenatal growth curve and the age at which growth is interrupted by birth (G). It is assumed that the rate of growth (R) immediately before birth is a sufficient index of the growth function and that the rate of growth is uniform at this time to a sufficient degree of approximation. In substituting gestation period for interval a small correction is desirable. On grounds which need not be gone into here, it is estimated that the correlation between interval and true gestation period is about .95. No correction is necessary for birth weight since there is little or no growth in the first day after birth. The correlations involving interval must be divided by .95 to obtain estimates of those involving gestation period.

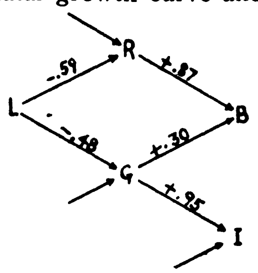


Fig. 12

$r_{BG} = +.56$, $r_{GL} = -.48$, while $r_{BL} = -.66$ is unchanged.

Minot's problem resolves mathematically the analysis of the observed correlation between birth weight and size of litter into the sum of two composite path coefficients representing the two postulated paths of influence.

$$r_{BL} = P_{BRL} + P_{BGL}$$

The method furnishes at once four equations for determining the values of the four path coefficients. One of these expresses the complete determination of B by R and G. The others are the expressions for the three known correlations.

(35) $$P_{BR}^2 + P_{BG}^2 + 2\,P_{BR} \cdot P_{BG} \cdot P_{RL} \cdot P_{GL} = 1$$

(36) $$P_{BR} \cdot P_{RL} + P_{BG} \cdot P_{GL} = -.66$$

(37) $$P_{BR} \cdot P_{RL} \cdot P_{GL} + P_{BG} = +.56$$

(38) $$P_{GL} = -.48$$

These are not all linear equations, a condition which generally distinguishes this sort of applicaton of the method from the calculation of partial regression coefficients. In the present case, however, there is no difficulty in the solution.

$P_{BR} = +.87,\quad P_{BG} = +.30,\quad P_{RL} = -.59,\quad P_{GL} = -.48,$

$P_{BRL} = -.51,\qquad P_{BGL} = -.15,\qquad r_{BL} = -.66.$

The result is an analysis of the correlation between birth weight and size of litter into two components whose magnitudes indicate that size of litter has more than three times as much linear effect on birth weight through the mediation of its effect on growth as through its effect on the length of the gestation period, contrary to the results of Minot's verbal analysis.

In this case, the answer to Minot's question might have been obtained from a set up mathematically identical with that used in multiple regression (after correcting the correlations with interval to obtain estimates of those with true gestation period.)
By equation 24,

$$(39)\qquad r_{BL} = P_{BL} + P_{BG} \cdot r_{GL}$$

$$(40)\qquad r_{BG} = P_{BL} \cdot r_{GL} + P_{BG}\,.$$

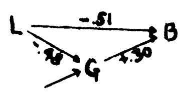


Fig. 13

The term $P_{BL} = -.51$ can be interpreted as measuring the influence of size of litter on birth weight in all other ways than through gestation period. In other cases, however, proper causal analysis may require a set up utterly different from that used in obtaining the best estimation equation. There is no routine method of making the proper diagram in the former case. This seems to have occasioned more misunderstanding than anything else among those who have attempted to apply the method. One author in a critique of the method, took the form of diagram intended to represent the sequential relations in the case of guinea pig weight and arranged some variables relating to basal metabolism in man in the same scheme in an arbitrary

way and then complained of the meaningless and absurd results which he obtained!

Transpiration of Plants

The contrast between the kind of set up appropriate to an estimation equation and that for evaluation of a causal interpretation was illustrated early (1921a) in connection with a study of the data of Briggs and Shantz on transpiration in plants. The reader is referred to the paper for the details, but it may be appropriate here to compare the different diagrams used. The authors obtained the total daily transpiration of a number of plants. The environmental factors studied were total solar radiation (R), wind velocity (W), air temperature in the shade (T), rate of evaporation from a shallow tank (E), and wet bulb depression, sheltered from sun, but not wind (B). To avoid seasonal effects, the logarithms of ratios for successive days were used instead of absolute values.

An estimation equation for wet bulb depression was obtained in terms of wind velocity, solar radiation and temperature (Fig. 14).

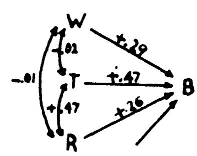


FIG. 14

It was pointed out that for causal analysis, radiation should be omitted as not affecting wet bulb depression in the shade, while a factor not directly measured, absolute humidity (H) should be included. There should be complete determination of B by W, T and H. As so arranged, there are two more unknown coefficients than known ones. It was assumed that there was no correlation between absolute humidity and wind velocity. The necessary additional equation was obtained from the theoretical multiple regression equation relating B to W, T and H, by substituting the extreme differences in wet bulb depression, temperature and wind velocity of the average daily cycle and assuming the absence of any such cycle in absolute humidity. Possibly this was not wholly justified in this case. If so, no numerical evaluation of the chosen point of view could be made.

Even in such cases, the attempt at analysis by path coefficients may be valuable in locating deficiencies in the data already collected and suggesting the kinds of new data which should be obtained.

The final set up used in relating transpiration T_r and evaporation from a tank to wet bulb depression and the chosen environmental factors is given in figure 15 with the values of the path coefficients and correlations. Determinations were made for 10 varieties of plants. These gave fairly consistent results which are averaged in Fig. 15 although there were certain interesting differences. There was a marked difference between the transpiration of the plants and the rate of evaporation from the tank in the relative importance of the various factors.

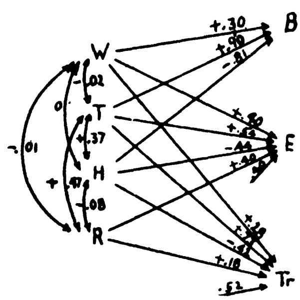


FIG. 15

The Relative Importance of Heredity and Environment

Among the most satisfactory applications to causal relations are to problems of genetic determination. The development of an organism is the product of the confluence and interaction of two distinct streams of causation, heredity and environment. The interaction between the hereditary influences emanating from the nuclei of the cells of the organism and the influences coming from outside these cells, but largely from other parts of the body, where they in turn are the products of heredity and cell environment and so on back to the one cell stage are complex enough, but if we go back of this to the ultimate factors: the array of genes assembled at fertilization and the environmental conditions external to the organism, the sequential relations are for the most part clear. The problem is that of determining the relative importance of *differences* in heredity and of *differences* in environment in determining differences in the characteristics of individ-

uals in a given population. The principal complications are the possibilities of nonlinearity in the combination effects of different genes, of different environmental factors, and of heredity and environmental factors in relation to each other.

We will review a case, the amount of white in the coat pattern of certain strains of guinea pigs, in which such combination effects appear to have been of negligible importance (Wright 1920, 1926b). A stock of guinea pigs was maintained for many years by the U. S. Bureau of Animal Industry without outcross, but with the avoidance of even second cousin mating. The correlation between mated individuals (143 pairs) was $+.060 \pm .083$, indicating that mating actually was at random in respect to coat pattern. The correlaton between parent and offspring averaged $+.191 \pm .024$ [8] with no significant differences in relation to sex.

By the theory developed on page 167, $r_{PO} = ab$. Allowing for incomplete determination by heredity this becomes

(41) $$r_{PO} = .19 = h^2ab = \tfrac{1}{2}h^2.$$

Thus $h^2 = .38$ leaving .62 for determination by environment. The correlation between litter mates averaged $+.282 \pm .027$. In the case of litter mates it is necessary to distinguish two groups of environmental factors — ones common to litter mates (E) and ones peculiar to individuals (D). From the diagram

(42) $$r_{O_1O_2} = 2h^2a^2b^2 + e^2 = \tfrac{1}{2}h^2 + e^2$$

where e^2 $(= p^2_{OE})$ is the determination by common environment. Its value is .09, leaving $d^2 = .53$ as the determination by nongenetic factors not common to litter mates. It seems rather surprising that the environment common to litter mates should determine so little in a character

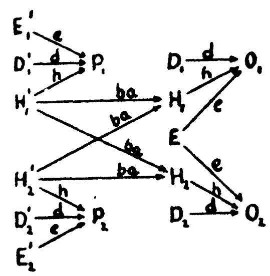


FIG. 16

[8] These standard errors, obtained from values in different subdivisions of the data are larger than would be obtained from the 3881 parent-offspring pairs, which however necessarily involve much repetition of individuals.

graded at birth, but only very minor effects of this sort have been discovered experimentally, the most important (contributing .036 to the correlation of litter mates) being an effect of the age of the mother. The high degree of asymmetry of the pattern in individual animals is in harmony with a large element of chance (somatic mutation?) in the determination of pigmented areas.

The above estimates ($h^2 = .38$, $d^2 = .53$, $e^2 = .09$) are estimates of the portion of the variance due to heredity, non-genetic factors peculiar to individuals, and common environment, respectively. They are the portion of the variance which should be eliminated by control of each factor. It is not possible to control the rather intangible environmental factors but hereditary *variation* can be eliminated by close inbreeding (decrease of heterozygosis being about 19% per generation under brother-sister mating). It happened that a number of piebald stocks were on hand, each descended from a single mating after several generations of inbreeding. These differed markedly in average percentage of white in the coat, although individuals of each varied widely about their family averages. Crosses between strains at opposite extremes gave intermediate offspring, justifying the assumption of no dominance. The family (No. 35) most advanced in inbreeding was descended from a single mating in the 12th generation of brother-sister mating, but even in it there was variation from nearly solid color to solid white. As expected by theory, very little, if any, of this variability was hereditary. The correlation between parent and offspring was only $+.024 \pm .020$. The correlation between litter mates was $+.103 \pm .025$, again indicating only a small amount of influence of environment common to litter mates.

The standard deviation, measured on an appropriate scale[4]

[4] On a percentage scale of measurement, necessarily limited at 0% and 100%, a given factor has more effect near the middle of the range than near the limits. The appropriate transformation of the scale X, ranging from 0 to 1 is $X' = prf^{-1}(x-50)$, where prf^{-1} is the inverse probability function, the direct function being defined in the form $prf\, x' = \frac{1}{\sqrt{2\pi}} \int_0^{x'} e^{-\frac{z^2}{2}} dz$. (Wright 1926a).

came out .574 (about 22% of the area of coat in the neighborhood of 50%). In the random bred stock the standard deviation was 0.782 (about 28%). The variance of the stock in which hereditary variation had been eliminated was thus $54\% = \left(\frac{.574}{.782}\right)^2$ of that of the random bred stock. This agrees as well as could be expected with the estimate of 62% of the variance of the latter as nongenetic, based on the parent offspring correlation, although not as well as an earlier estimate made when the numbers were smaller (nongenetic variance 58% as deduced for a parent-offspring correlation of $+.21$ in random stock, variance of inbred family 57% of that of random stock, *i.e.* $\frac{.364}{.643}$).

Case of Human Intelligence

Another illustration of the difference between a quantitative interpretation and a multiple regression formula has been given (Wright 1931a) using data of Miss B. S. Burks, on the roles of heredity and environment in determining human intelligence. These data consisted of intelligence tests of 104 California children, tests of their parents and in addition grades of home environment. Similar data were obtained of 206 children adopted at an average age of 3 months, and of their foster parents and home environments. The correlations as used were corrected by Miss Burks for attenuation.

If the purpose is to obtain the best estimation for children in terms of their parents and environments, the variables are to be related as in figure 17 in which C is child's intelligence, P is midparent and E is the measure of home environment.

Normal Equations		Children (Own)	Children (Adopted)
(43)	$p_{CE}\cdot r_{EP} + p_{CP} = r_{CP}$	$= +.61$	$+.23$
(44)	$p_{CE} + p_{CP}\cdot r_{EP} = r_{CE}$	$= +.49$	$+.29$
(45)	r_{EP}	$= +.86$	$+.86$
Solutions:	p_{CP}	$= +.72$	$-.07$
	p_{CE}	$= -.13$	$+.35$

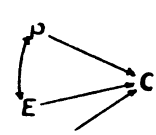


FIG. 17

The solutions of the normal equations in the two bodies of data give what at first sight appear to be contradictory results. There is no apparent reason why environment should not play as great a role in shaping intelligence in one case as in the other, yet it turns out that while the partial regression of child's IQ on home environment is significantly positive in the foster data, it is negative as far as it goes, in the case of own children.

The point that is sometimes overlooked is that the arrangement for obtaining the best possible prediction equation does not necessarily yield coefficients which have any simple interpretation. This is obviously the case here. If child's IQ is affected both by heredity and environment, the same is presumably true of parental IQ. In so far as the latter is determined by environment it is not a causal factor in relation to child's *heredity*. A diagram intended to represent causal relation *must* represent parental IQ as merely correlated (two headed arrows) with child's heredity and child's environment. Another complication which must be represented is the correlation of heredity with environment. Good heredity in a family will tend to create a good environment and vice versa. The simplest possible *interpretative* diagram for own children is thus of the type of figure 18. That for foster children is given in figure 19.

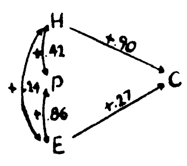


FIG. 18

Even these are doubtless too simple since heredity is represented as the only factor apart from the measured environment. Any estimates of the importance of hereditary variation will thus be maximum.

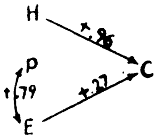


FIG. 19

The two correlations given by Miss Burks in the case of the foster data (Fig. 19)

$$(r_{CE} = +.29, \quad r_{CP} = .23)$$

yield the value $r_{EP} = +.79$

for the correlation between home environment and midparental IQ. The actual correlation was not published for the foster data, but there is no reason why it

should differ significantly from that in the other data in which the value was $+.86$. There is reasonable agreement.

In the case of own children, three correlations of interest here, were published, $r_{CE} = +.49$, $r_{CP} = +.61$, $r_{EP} = +.86$. But this is not enough to give a solution for the coefficients of the 5 indicated paths. The assumption of complete determination of C by H and E gives a fourth equation, still an inadequate number. No solution is possible, a situation which as previously noted, very frequently arises in such analysis, even when one makes the most simplified possible qualitative representation of the causal relations. A great deal of utterly unwarranted verbal interpretation of correlation coefficients would be avoided if the authors took the trouble to represent their ideas in diagrammatic form and noted whether or not the number of equations possible from the data (known correlation coefficients and known cases of complete determination) was as great as the number of paths in this diagram.

In the present case, another equation can be obtained by borrowing from the foster data. Environment should make approximately the same contribution to IQ in both groups of children. The concrete partial regression coefficients

$$C_{CH} \quad \left(= P_{CH}\frac{\sigma_C}{\sigma_H}\right) \quad \text{and} \quad C_{CE} \quad \left(= P_{CE}\frac{\sigma_C}{\sigma_E}\right)$$

should thus be approximately the same in the foster as in the own children. Assuming that σ_H and σ_E are the same in both cases, the ratio $\frac{P_{CE}}{P_{CH}}$ from the foster data may be accepted for the group of own children. The five equations now available are as follows:

	Equations	*Solution*
(46)	$r_{EP} = +.86$	
(47)	$r_{CE} = +.49 = P_{CE} + P_{CH} \cdot r_{HE}$	$P_{CE} = +.27$
(48)	$r_{CP} = +.61 = P_{CE} \cdot r_{EP} + P_{CH} \cdot r_{HP}$	$P_{CH} = +.90$
(49)	$P_{CE} = +.302\, P_{CH}$	$r_{PH} = +.42$
(50)	$P_{CE}^2 + P_{CH}^2 + 2P_{CE} \cdot P_{CH} \cdot r_{HE} = 1$	$r_{HE} = +.24$

The solution assigns reasonable values in all cases and shows that there was no real disagreement involved in the relation of the two groups of children to their environments.

It was noted that this analysis gives a maximum estimation of the role of heredity. An attempt was made to obtain a minimum estimate compatible with acceptance of the observed correlations, by carrying the analysis back a generation and assuming as much similarity in the determining factors of successive generations as the data permit.

Such analysis requires separate treatment of heredity (H) as a factor of development, and heredity or genotype (G) as the linear system of gene effects which best approximates the former. Departures from linearity in the effects of allelomorphs (dominance) and in the effects of nonallelomorphs (epistasis) are common. Moreover there may be non-linearity in the combination effects of heredity and environment. Thus a certain genetic complex in the guinea pig $(c^d c^a BB)$ produces more melanin pigment at low temperatures than does a certain other $(CCbb)$ but less at high temperatures (Wright 1927). The subject is too involved for detailed discussion here but it may be noted that in general correlations between deviations due to dominance and epistasis must be taken account of.

In the case of Miss Burk's data, there is no possible way of distinguishing the effects of environmental factors not included in the measurement of home environment from the contributions of dominance and epistasis or from non-linearity in the combination effects of heredity and environment. In the attempt at obtaining a minimum estimate of heredity, these three very diverse factors were put together in a miscellaneous group M. The diagram of relation used is given in Fig. 20. Child's genotype (G) is represented as partially determined by midparental genotype (G'), the residual variability being that of Mece -

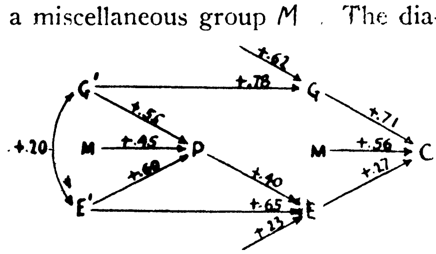


FIG. 20

lian segregation. Child's environment is treated as in part determined directly by midparental intelligence P and in part tracing to the environment of the preceding generation, E'.

The path coefficient relating genotype of midparent to that of child could be estimated, assuming Mendelian heredity and taking into account a correlation of $+.70$ between father and mother. It turned out to be mathematically impossible to assign the same values to the path coefficients of the parental generation as in the offspring generation, but this is not surprising since the parents were tested as adults instead of young children. The solution for the parent generation was to some extent indeterminate but within rather narrow limits, on making what seemed the most reasonable assumptions. The values reached are given in figure 20. The path coefficient for influence of hereditary variation lies between the limits $+.71$ (if dominance and epistasis are lacking) and $+.90$.

Analysis of Size Factors

The first published application of the method was to the interpretation of a system of correlations of bone measurements (length and breadth of skull, lengths of humerus, femur and tibia) in a population of rabbits (Wright 1918). The 10 observed correlations were accounted for primarily as due to a single general factor (not necessarily acting proportionately on the 5 variates). The residuals which appeared were attributed to group factors.

In a recent paper (1932b) the same figures, two other sets of figures for rabbit populations (F_1 and F_2 of a wide cross) and figures from a flock of hens have been analyzed by a somewhat improved method. A set of n variables yields $\frac{n(n-1)}{2}$ correlation coefficients and hence the same number of observation equations of the type $r_{ab} = p_{AG} \cdot p_{BG}$, where A and B are two of the variables and G is the general factor and it is assumed for the moment that the correlations are due solely to differences in general size. The residuals are minimized by the method of least squares.

This method necessarily gives residuals which are as likely to be negative as positive. The interpretation is more satisfactory if the path coefficients relating each measurement to the general factor are all reduced by the proportion necessary to eliminate significant negative residuals. It happened that in each of the 4 sets of data studied, the most important negative residuals were those between the skull and hind leg measurements, and the method followed was to eliminate the average of these.

The important positive residuals in all cases indicated natural group factors—a head group, a general leg group, a foreleg group (in the one case in which both humerus and ulna were measured) and a hind leg group. Other indications such as a slightly closer relation between head and foreleg than between head and hind leg, slightly closer relation between proximal leg bones (humerus and femur) than between non-homologous and hind leg bones (humerus and tibia) were less certain. Figure 21 shows the system of path coefficients arrived at in the case of the fowl measurements. The squares of these give the degree of determination in each case by the general factor, the group factors and special factors.

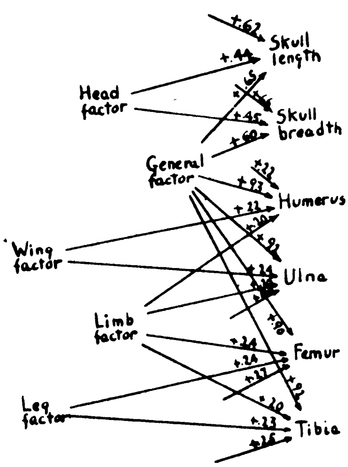


FIG. 21

The Use of Partial Correlation in Interpretation

Partial correlation coefficients have sometimes been used in the attempt to interpret systems of correlated variables apparently on the theory that the reduction or elimination of a correlation between two variables on holding a third constant demonstrates the latter to be causally responsible for the correlation. The

method at first sight seems analogous to that of the experimentalist in attempting to control all sources of variation except those in which he is interested. This, however, is a delusion in the case of correlation (as opposed to regression) coefficients (Wright 1921a) and the method of path coefficients was developed because of the unsatisfactory nature of interpretation based on partial correlation. As R. A. Fisher (1925) has stated, "In no case, however, can we judge whether or not it is profitable to eliminate a certain variable unless we know or are willing to assume a qualitative scheme of causation."

This point can be illustrated by considering a system of 3 variables, A, B, and C in which the following correlations have been found.

$$r_{AB} = .50 \qquad r_{BC} = .50 \qquad r_{AC} = .25.$$

By substitution in the usual formula, $r_{AC \cdot B} = 0$. This is compatible with the interpretation, represented in figure 22, that B is an intermediary in a single chain of causation connecting C and A.

$$r_{AC} = P_{AB} \cdot P_{BC} = r_{AB} \cdot r_{BC} = .25.$$

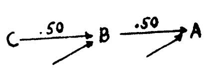


FIG. 22

Another interpretation is that B is the only common factor

$$r_{AC} = P_{AB} \cdot P_{CB} = r_{AB} \cdot r_{BC} = .25 .$$

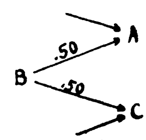


FIG. 23

But it is also possible that B may be the product of the interaction of two correlated factors A and C

$$r_{AB} = P_{BA} + P_{BC} \cdot r_{AC} = .50 .$$

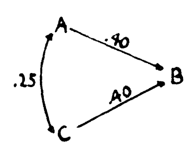


FIG. 24

Finally A, B and C may be correlated with each other through reciprocal interactions or through complexes of unknown common factors, making impossible anything beyond the mere descriptive use of the correlation coefficients, or the calculation of estimation equations.

The first step in the application of the method of path coefficients is to bring clearly into the open the system of functional relations among the variables which seems significant for purposes of interpretation. In the majority of cases, verbal interpretations which seem reasonable enough as long as the basic postulates are kept discretely in the subconscious mind become obviously crude and inadequate when expressed in a diagram. Occasionally, however, statistical systems are capable of some interpretation.

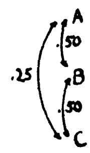


FIG. 25

Difficulties in Causal Analysis

There are a great many systems of correlated variables for which no interpretation can be suggested in terms of sequential relations. Among these are cases in which there is prevailingly mutual interaction between the variables instead of action in one direction. The branches of science differ considerably in the type of relation which predominates.

As already noted the developmental process of organisms is essentially a one way process, and the ultimate factors of development, heredity and environment act on it without being acted upon. A method of analysis which takes account of the sequential relations is thus imperatively called for in genetics.

A case in which such analysis would not be possible may be illustrated by the relations among the various properties of the blood, as discussed by L. J. Henderson. The physiological mechanisms are such that alteration of any one brings about immediate readjustments in the values of the others. What one wishes to determine are the functional relations, whether in the form of equations or of nomograms. If such a system were studied by correlational methods the best that could be done would be to attempt to approximate the functional relations by multiple regression (linear or curvilinear as the case required).

There is usually rapid reciprocal action among the variables of interest to the economist or sociologist and the correlations

among the simultaneous deviations cannot, in most cases, be treated as due to lines of one way causation among these variables themselves. Thus the price of a commodity cannot properly be treated as caused by the amount marketed or vice versa. The exception is where one variable is clearly external to the social system in question as is the influence of weather on crop yield.

There is more likelihood of being able to represent the various simultaneous deviations as direct consequences of the system of deviations of the preceding year (together with the clearly external contemporary factors) but even here, a causal diagram can be set up only after a most careful consideration of the realities of the case. There may be lags of greater duration than one year and a correlation between two variables in successive years may trace to more remote common factors rather than to a direct line of causation from the earlier to the later.

Corn and Hog Correlations

These points were illustrated by a study of corn and hog correlations (Wright 1924). An attempt was made to analyze the play of interacting factors responsible for the annual fluctuations from the general trends in production and price of hogs during the relatively undisturbed period between the Civil War and the World War. It was shown that variation in the corn crop and certain interrelations among the hog variables themselves determined from 75 to 85% of the variance of the latter. The annual fluctuations about the trend during the period of years from 1871 to 1915 inclusive (so far as data were available) were found for corn acreage, yield, crop and price and for western and eastern wholesale hog packs and for farm price of hogs. The fluctuations were found separately for the summer and winter seasons for western wholesale hog pack and the corresponding live weight, pork production (product of preceding) and price. Correlation coefficients were found not only for the same year but between variables separated by one, two and often three years. Altogether 510 correlation coefficients were calculated.

Most of these coefficients could be given reasonable enough verbal interpretations, but there was no assurance that the "obvious" interpretation in one case, was compatible with an equally "obvious" interpretation in another. The problem was to represent all of these verbal interpretations in a single diagram and determine path coefficients which would account simultaneously for the entire system of correlation coefficients. With 510 correlation coefficients and 4 cases of complete determination, one could write 514 simultaneous equations to determine the values of whatever system of path coefficients had been used. Theoretically one could introduce the same number of different paths into the diagram. It would not be practicable, however, to deal with such a large number of unknown quantities and even if practicable, the complexity of the system would defeat the purpose of the analysis. The problem thus resolved into the discovery of a simple system of relations which would give a reasonably close approximation to all of the correlation coefficients.

It has been emphasized that the method of path coefficients is not intended to accomplish the impossible task of deducing causal relations from the values of the correlation coefficients. It is intended to combine the quantitative information given by the correlations with such a qualitative information as may be at hand on causal relations to give a quantitative interpretation. The analysis of cases such as the present and that preceding (size factors), in which the equations far outnumber the coefficients to be determined, may appear to be exceptions to this statement, but even here only such paths are tried which are appropriate in direction in time and which can be given a rational interpretation.

Considerable experimentation was necessary before a simple system could be found which gave even moderately satisfactory results. The procedure followed was to list the highest five correlations of each variable with a preceding variable. It turned out that the corn variables were so nearly independent of conditions in preceding years that they might be treated practically as

independent in relation to the hog situation. The variations in corn crop depended largely on variations in yield $(p_{CY} = +.90)$ and secondarily on variations in acreage $(p_{CA} = +.45)$ Corn price showed a correlation of $-.80$ with the crop.

Among the hog variables, the maximum correlations were with those which indicated most directly the amount of breeding (average summer weight (SW) of the same year, winter pack (WP) a year and a half later, between which there was a correlation of $+.78$), and with the preceding prices of corn and of hogs. The four variables: breeding (B), summer (S) and winter (W) price of hogs and price of corn (P) were thus chosen as a central system. 36 equations could be written involving these (using jointly the two indicators of breeding). Values of 13 path coefficients were tested by repeated trial and error until it seemed that no change (of the order of .05) would give improvement. The system reached is shown in figure 26 in which primes refer to preceding years.

The other variables were then appended to this system, also by the trial and error method. Corn crop was used in place of corn price, however. The results are shown in figures 27 and 28. These bring out the very different characteristics of the summer pack (SP) (consisting of a very heterogenous lot of hogs) and the winter pack, (WP) largely consisting of the spring pig crop. Average summer and winter live weights are represented by (SW) and (WW) in figure 27.

The general conclusions were that the dominating features of the hog situation are the corn crop and its price, and an innate tendency to fall into a cycle of successive overproduction and underproduction, two years from one extreme to the other, depending mainly on two compound paths:

$$p_{BWB''} = -.42 \text{ and } p_{BSB''} = -.14.$$

The 32 indicated path coefficients together with 10 others relating total annual western pack to its components, eastern pack to western pack and prices, and farm price to packer's price, ac-

counted for the 510 observed correlation coefficients with an average error of only .09 neglecting sign. The most serious discrepancies were in certain correlations involving corn acreage and yield which were intentionally ignored for the sake of avoiding complexity in the relations of the more important variables.

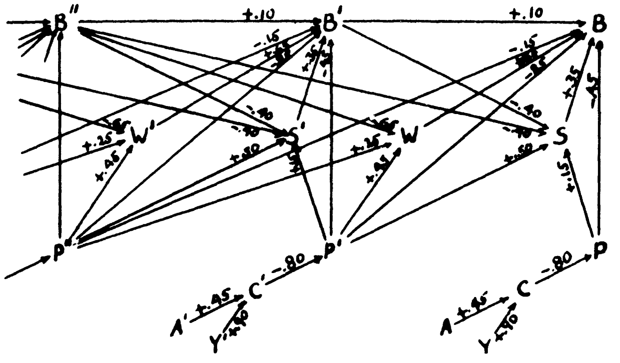


FIG. 26

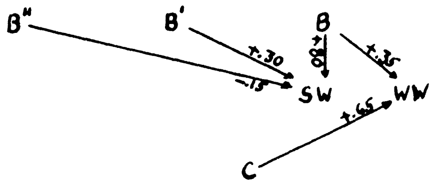


FIG. 27

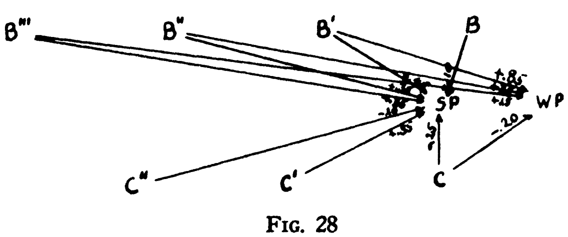


FIG. 28

The Elasticities of Supply and Demand

In the preceding illustration, market supplies, prices, etc. were related to preceding conditions largely by a trial and error process of finding the system which would work best and without much

regard for theoretical considerations. In the following more theoretical approach I have collaborated with Dr. P. G. Wright. The purpose is to interpret observed series of prices and quantities marketed as functions of two hypothetical variables, the conditions of supply and demand. Only a brief reference has previously been published (P. G. Wright 1928).

The demand for a given commodity and given market is treated as that function of all economic factors (prices, wages, etc.) which determines the quantity which would be purchased under any set of postulated conditions. The supply function, similarly, is treated as that function of all economic factors (prices, manufacturing costs, weather, etc.) which determines the quantity which would be offered for sale under any set of postulated conditions. The actual values which these functions take at a given moment tend to be the same, the price of the commodity itself being the immediate factor which shifts to such a value as to make them identical.

We shall deal with the annual percentage deviations in quantities and prices, whether from the preceding year or from the estimated trend of a series of years, instead of absolute values. The relative merits of these two procedures need not be gone into

Let X represent values on a scale of *percentage* change in quantity and Y values on a scale of *percentage* change in the price of the commodity in question. Let Z_1, Z_2, etc. represent other economic factors of demand or supply or both on whatever scales are most suitable. The demand and supply functions themselves as percentage deviations in quantities under postulated conditions may be represented by X_d and X_s respectively.[5]

$$X_d = f_d\,(Y_1,\ Z_1,\ Z_2, \cdots,\ Z_d, \cdots) \tag{51}$$

$$X_s = f_s\,(Y_1,\ Z_1,\ Z_2, \cdots,\ Z_d, \cdots)\,. \tag{52}$$

[5] If the absolute quantities are represented by U and the absolute prices by V, $X = \frac{\Delta U}{U}$ and $Y = \frac{\Delta V}{V}$. It is customary to define the demand and supply functions in terms of the absolute values, but for the present purpose it is more convenient to define them in relation to the percentage deviation.

Assume that these functions are of such a nature that the deviations in price can be separated linearly from the other factors to a sufficient degree of approximation. This does not imply lack of correlation between price and the others.

(53) $X_d = \eta Y + D$ where $D = f_d'(z_1, z_2, \cdots, z_d, \cdots)$

(54) $X_s = e Y + S$ where $S = f_s'(z_1, z_2, \cdots, z_s, \cdots)$.

The demand function is here analyzed into two variable components, a multiple of the price deviation (ηY) and the deviation (D) in the quantity which would be purchased if there were *no price deviation* $(Y=0)$. The supply function is similarly analyzed into a different multiple of the price deviation $(e Y)$ and the deviation (S) in the quantity which would be offered for sale in the absence of a price deviation. Thus D and S measure the strength of demand and supply apart from price and will be spoken of as measures of demand and supply. For given values of D and S the equations define two straight lines which describe the momentary demand and supply situations respectively (Fig. 29). Their slopes relative to the Y-axis are given by η and e respectively. These slopes are in accordance with the customary definitions of the elasticities of demand and supply, recalling that X and Y are percentage deviations.[6]

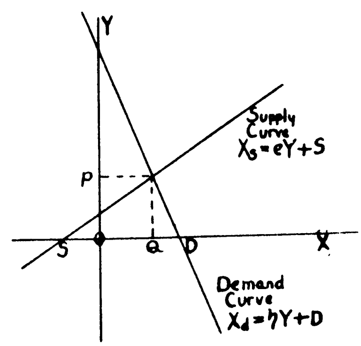


FIG. 29

According to the usual theory, the actual quantity which changes hands and the actual price are determined by the point of intersection of the supply and demand curves. Under the approximations previously assumed, and assuming constancy of the elasticities, but variation of D and S, the percentage deviations

[6] The ratio $\frac{X}{Y} = \frac{\Delta u/u}{\Delta v/v}$ where U and V are absolute quantities and prices respectively. The ratio $\frac{X_S}{Y}$ is the elasticity of supply (e) if $S = 0$, and $\frac{X_D}{Y}$ is the elasticity of demand (η) if $D = 0$.

in quantity (Q) and in price (P) are linear functions of D and S. Their values may be represented as determined by multiple regression equations. It will be convenient to use single letters for the path coefficients.

$$P = p_1 \frac{\sigma_P}{\sigma_D} D + p_2 \frac{\sigma_P}{\sigma_S} S \tag{55}$$

$$Q = q_1 \frac{\sigma_Q}{\sigma_D} D + q_2 \frac{\sigma_Q}{\sigma_S} S. \tag{56}$$

The elasticity of supply may be obtained from the ratio of Q to P under a fixed average supply situation ($S = 0$) but variable demand.

$$e = \frac{q_1 \sigma_Q}{p_1 \sigma_P} \tag{57}$$

Similarly, elasticity of demand is given by the ratio of Q to P when D equals zero.

$$\eta = \frac{q_2 \sigma_Q}{p_2 \sigma_P}. \tag{58}$$

Since the standard deviations are obtainable directly from the data it is merely necessary to find the values of the path coefficients in order to calculate the two elasticities.

A diagram can be set up as in Fig. 30 indicating primarily that P and Q are different linear functions of D and S. Three equations can be written at once; two indicating complete determination of P and Q by D and S, and one representing the correlation between P and Q.

$$p_1^2 + p_2^2 + 2 p_1 p_2 r_{SD} = 1 \tag{59}$$

$$q_1^2 + q_2^2 + 2 q_1 q_2 r_{SD} = 1 \tag{60}$$

$$p_1 q_1 + p_2 q_2 + (p_1 q_2 + p_2 q_1) r_{SD} = r_{QP}. \tag{61}$$

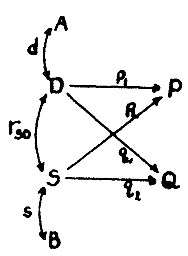


FIG. 30

Unfortunately these three equations involve 5 unknowns. Other data must be brought to bear on the problem before any solution is possible.

The diagram suggests two possible sources of

additional data. If any measurable quantity (A) can be found which is correlated with the demand situation but which can safely be assumed to be independent of the supply situation, $(r_{AS} = o)$ we can write two new equations representing the correlations r_{AP} and r_{AQ} respectively at the expense of only one additional unknown $(r_{AD} = d)$. We have now 5 equations and 6 unknowns. If it can safely be assumed that there is no correlation between the demand and supply situations $(r_{SD} = o)$, a solution is possible. If such an assumption with regard to r_{SD} does not seem justified, it may be possible to find a quantity (B) correlated with the supply situation (as measured by S but of such a nature that no correlation with the demand situation need be postulated. The correlation r_{BP} and r_{BQ} make possible two more equations, with only one more unknown $(r_{SB} = s)$ bringing the number of equations and unknowns both up to 7. The path coefficients and hence the elasticities are now determinate. The additional equations are as follows:

(62) $r_{AP} = p_1 d$ (64) $r_{BP} = p_2 s$

(63) $r_{AQ} = q_1 d$ (65) $r_{BQ} = q_2 s$.

The hog and corn data referred to in the preceding section were not obtained with the present purpose in mind, but may furnish rough illustrations of the method. The total weight of hogs, marketed at the principal markets in the summer season (March to October) 1889–1914, and the reported price may be considered first. Absolute instead of percentage deviations from trend were used but the correlations should not be affected much and coefficients of variation may be used in place of the standard deviations on a percentage scale. The most important single factor affecting the summer hog pack was shown to be the corn crop of the preceding year. It is assumed that it is a factor of type B, correlated with the supply situation as measured by S but not with the demand for pork as measured by D. It is further as-

sumed that there was no correlation between the supply and demand situations.

Data

Coefficient of variation—Price	$\sigma_P = 15.86$
Quantity	$\sigma_Q = 10.89$
Correlation—Price with quantity	$r_P = -.63$
Correlation—Hog price with preceding corn crop	$r_{P,1} = -.47$
Correlation—Weight of pack with preceding corn crop	$r_Q = +.64$

Equations		*Solution*	
$p_1^2 + p_2^2 = 1$		$p_1 = +.686$	$e = +.133$
$q_1^2 + q_2^2 = 1$		$p_2 = -.728$	$\eta = -.944$
$p_1 q_1 + p_2 q_2 = -.63$		$q_1 = +.132$	
$p_2 s = -.47$		$q_2 = +.991$	
$q_2 s = .64$		$s = +.646$	

The solution indicates very little elasticity of supply ($e = +.133$) but a very considerable elasticity of demand ($\eta = -.944$).

Similar data were given for the winter weight of pack (1870–1914). The largest correlation with a factor of preceding years was with average summer live weight of hogs, one and one half years before. This factor (an index of amount of breeding) is again assumed to be related to the supply but not to the demand situation, and again it is assumed that the supply and demand situations vary independently of each other.

Data	*Solution*	
$\sigma_P = 12.59$	$p_1 = +.656$	$e = +.110$
$\sigma_Q = 18.75$	$p_2 = -.755$	$\eta = -.884$
$r_{PQ} = -.68$	$q_1 = +.108$	
$r_{PB} = -.63$	$q_2 = +.994$	
$r_{QB} = +.83$	$s = +.835$	

The results are remarkably close to those of the quantity and price of summer pork. The low elasticities of supply are to be expected of an agricultural commodity the quantity of which is largely determined in advance and by factors independent of the market demand and which once produced must largely be marketed.

I am indebted to my colleague, Professor Henry Schultz, for data on the quantity and price of potatoes marketed annually from 1896 to 1914 and the suggestion that it would be interesting material for analysis by this method. Trends had been fitted by Professor Schultz and trend ratios of quantity and price obtained.

Data

Standard deviation of price ratios	σ_P	$= .185$
Standard deviation of quantity ratios	σ_Q	$= .130$
Correlations		
Price—quantity (same year)	r_{PQ}	$= -.852$
Price—quantity (preceding year)	$r_{PQ'}$	$= +.570$
Price—price (preceding year)	$r_{PP'}$	$= -.562$
Quantity—quantity (preceding year)	$r_{QQ'}$	$= -.522$
Quantity—price (preceding year)	$r_{QP'}$	$= +.651$.

It is assumed again that there is no correlation between supply and demands situations ($r_{SD} = 0$) and that the price (as a trend ratio) is a factor of type *B* affecting the supply of the following year but without influence on the demand of the following year. The solution is as follows:

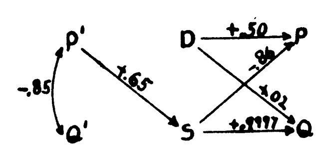


FIG. 31

$$p_1 = +.503 \qquad q_1 = +.024 \qquad e = +.034$$

$$p_2 = -.864 \qquad q_2 = +.9997 \qquad \eta = -.815$$

$$s = +.651$$

Figure 31 gives a graphical representation of the relation.

Again the virtual absence of elasticity of supply might perhaps have been anticipated. The size of crop is largely determined

before the price is known and the crop must be disposed of regardless of price. It is to be noted, however, that this result came out quite independently of any such assumption.[7] There are other checks on the theory. Two of the correlations reported above have not been used. According to the diagram of relations $r_{QQ'} = q_2 s r_{PQ} = -.554$. The observed value, $-.522$, is in good agreement. Also $r_{PQ'} = p_2 s r_{PQ} = +.479$.
The agreement with the observed value of $+.570$ is not as good as in the previous case, but considering the small number of years, is not bad.

The absence of elasticity of supply in the case of potatoes applies only within a single year. The fact that the supply is strongly correlated with the price of the preceding year $+.651$ indicates that in the long run there is considerable elasticity. The method of path coefficients readily lends itself to deduction of this long time elasticity.

Let $\bar{P}, \bar{Q}, \bar{A}$ and $\bar{B}$ be the hypothetical averages of P, Q, A and B respectively over an indefinite (n) period of years. The problem is to deduce the elasticities toward which the long time supply and demand curves tend, from knowledge merely of the correlations from year to year. The following equation can be written from figure 32, where a, b, c and g are path coefficients pertaining to the paths indicated.

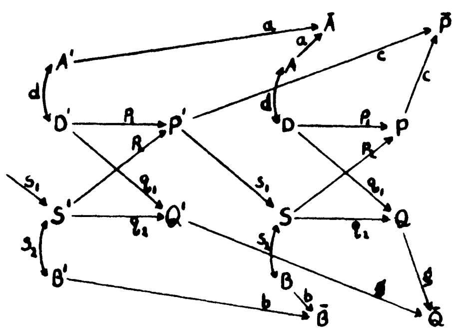


FIG. 32

[7] In two other cases studied by this method (P. G. Wright 1928) very different results were obtained. In the case of butter, the elasticity of supply came out 1.43, of demand —.62. In the case of flax seed, the elasticity of supply came out even greater, 2.39, while that of demand was —.80. But these are cases in which a high elasticity of supply is to be expected on a priori grounds. It is interesting to note that in cases in which it seems justifiable to assume a priori that there is no elasticity of supply $(e = o)$, it follows that $q_1 = o$, $q_2 = 1$, $p_2 = r_{QP}$ (still assuming $r_{SD} = o$) and finally that $\eta = \frac{\sigma_Q}{r_{PQ} \cdot \sigma_P} = \frac{1}{b_{PQ}}$.

$$(66)\quad r_{\bar{P}\bar{A}} = nc\, r_{P\bar{A}} = nc\,(p_1 r_{D\bar{A}} + p_2 r_{S\bar{A}})$$
$$= nc\,(p_1 r_{D\bar{A}} + p_2 s_1 r_{P\bar{A}}) = \frac{nc\, p_1 r_{D\bar{A}}}{1 - p_2 s_1}$$

$$(67)\quad r_{\bar{P}\bar{B}} = nc\, r_{P\bar{B}} = nc\,(p_2 r_{S\bar{B}})$$
$$= nc\, p_2 (s_1 r_{P\bar{B}} + s_2 b) = \frac{nc\, p_2 s_2 b}{1 - p_2 s_1}$$

$$(68)\quad r_{\bar{Q}\bar{A}} = ng\, r_{Q\bar{A}} = ng\,(q_1 r_{D\bar{A}} + q_2 r_{S\bar{A}}) = ng\,(q_1 r_{D\bar{A}} + q_2 s_1 r_{P\bar{A}})$$
$$= ng\left(q_1 r_{D\bar{A}} + \frac{q_2 s_1 p_1 r_{D\bar{A}}}{1 - p_2 s_1}\right) = ng\, r_{D\bar{A}}\left(\frac{q_1 + q_2 s_1 p_1 - q_1 p_2 s_1}{1 - p_2 s_1}\right)$$

$$(69)\quad r_{\bar{Q}\bar{B}} = ng\, r_{Q\bar{B}} = ng\, q_2 r_{S\bar{B}} = ng\, q_2 (s_1 r_{P\bar{B}} + s_2 b)$$
$$= ng\, q_2 (s_1 p_2 r_{S\bar{B}} + s_2 b) = \frac{ng\, q_2 s_2 b}{1 - p_2 s_1}$$

$$(70)\quad \bar{p} = \frac{\Sigma p}{n} = c\,\frac{\sigma_{\bar{p}}}{\sigma_p}\Sigma p \quad \therefore\ \sigma_{\bar{p}} = \frac{\sigma_p}{nc}$$

$$(71)\quad \bar{Q} = \frac{\Sigma q}{n} = g\,\frac{\sigma_{\bar{Q}}}{\sigma_Q}\Sigma Q \quad \therefore\ \sigma_{\bar{Q}} = \frac{\sigma_Q}{ng}.$$

Let η_L and e_L be the elasticities of long time demand and supply respectively.

$$(72)\quad \eta_L = \frac{r_{\bar{Q}\bar{B}}\,\sigma_{\bar{Q}}}{r_{\bar{p}\bar{B}}\,\sigma_{\bar{p}}} = \left(\frac{ng\, q_2 s_2 b}{1 - p_2 s_1}\right)\left(\frac{1 - p_2 s_1}{nc\, p_2 s_2 b}\right)\frac{\sigma_Q\, nc}{ng\, \sigma_p}$$
$$= \frac{q_2 \sigma_Q}{p_2 \sigma_p}$$

$$(73)\quad e_L = \frac{r_{\bar{Q}\bar{A}}\,\sigma_{\bar{Q}}}{r_{\bar{p}\bar{A}}\,\sigma_{\bar{p}}} = ng\, r_{D\bar{A}}\,\frac{q_1 + q_2 s_1 p_1 - q_1 p_2 s_1}{1 - p_2 s_1}\cdot\frac{1 - p_2 s_1}{nc\, p_1 r_{D\bar{A}}}\cdot\frac{\sigma_Q\, nc}{ng\, \sigma_p}$$
$$= \frac{q_1 + q_2 s_1 p_1 - q_1 p_2 s_1}{p_1}\cdot\frac{\sigma_Q}{\sigma_p} = e + s_1\left(q_2 - \frac{q_1 p_2}{p_1}\right)\frac{\sigma_Q}{\sigma_p}.$$

Thus a reaction of price of one year on the supply situation of the next does not tend to produce any difference between long time and short time elasticity of demand. It does make a difference, however, in long and short time elasticities of supply. In the case of potatoes substitution of values already found gives $e_L = +.52$ as the elasticity of the long time supply curve, insofar as determined by the reaction of the price of one year on the supply of the next. If it were legitimate to assume that there is no elasticity of supply within a year $(e = 0)$, the formula for e_L reduces to $s_1 \frac{\sigma_Q}{\sigma_p} = r_{QP'} \frac{\sigma_Q}{\sigma_p} = b_{QP'}$

Tests of Significance

In considering the reliability of path coefficients there are two questions which must be kept distinct. First is the adequacy of the qualitative scheme to which the path coefficients apply and second is the reliability of the coefficients, if one accepts the scheme as representing a valid point of view. The setting up of a qualitative scheme depends primarily on information outside of the numerical data and the judgment as to its validity must rest primarily on this outside information. One may determine from standard errors whether the observed correlations are compatible with the scheme and thus whether it is a possible one, but not whether it correctly represents the causal relation.

Having accepted a certain scheme with which the data are compatible, one would like to determine the reliability of the values reached for the path coefficients. Obviously no single formula can be given, applicable to all cases. The basic formulae of the method are ones for writing series of simultaneous equations, which must be solved to obtain the unknown path coefficients and correlation coefficients. These equations are in general non-linear with respect to the unknown quantities, making it impossible to express the solution in a general formula in which substitution can be made in routine fashion.

Certain principles can, however, be illustrated by the results

in simple cases. No attempt will be made here to deal with the complications due to small numbers. It will be assumed that the errors of sampling are in general so small in comparison with the values of the coefficients that second degree terms in the errors may be ignored. It is recognized that a more thorough treatment of the matter is much to be desired.

The simplest set up (Fig. 33) is that in which one variable V_0 is represented as a function of another V_1, and of a residual factor V_u. The equations are as follows:

(74) $p_{01} = r_{01}$

(75) $p_{01}^2 + p_{0u}^2 = 1$. From (74),

(76) $\sigma_{p_{01}}^2 = \sigma_{r_{01}}^2 = \dfrac{1 - r_{01}^2}{N}$.

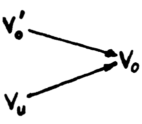


FIG. 33

From (75)

(77) $2\,p_{01}\,\delta p_{01} + 2\,p_{0u}\,\delta p_{0u} \doteq 0$,

(assuming as noted above that δp_{01} and δp_{0u} are small compared with p_{01} and p_{0u}).

(78) $\delta p_{0u} = -\dfrac{p_{01}}{p_{0u}}\,\delta p_{01}$

(79) $\sigma_{p_{0u}}^2 = \dfrac{p_{01}^2}{p_{0u}^2}\,\sigma_{p_{01}}^2 = \dfrac{r_{01}^2\,(1 - r_{01}^2)}{N}$

The standard error of the residual path coefficient in a system in which one variable is represented as determined by a number of others (Fig. 34) may be derived similarly

(80) $\sigma_{p_{0u}}^2 = \frac{1}{N}\, r_{0(12\cdots m)}^2 \left[1 - r_{0(12\cdots m)}^2\right]$.

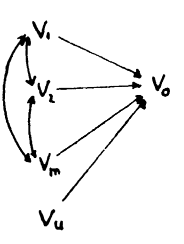
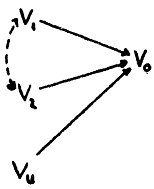


FIG. 34

Consider next the case in which variable V_0 is a function of two uncorrelated variables V_1 and V_2, and of residual factor V_u.

Two different solutions are obtained for $\sigma^2_{p_{01}}$ depending on the point of view. If it is accepted that V_1 and V_2 are wholly independent, except for the accidents of sampling, we have



(81) $$p_{01} = r_{01}$$

(82) $$\sigma^2_{p_{01}} = \sigma^2_{r_{01}} = \frac{(1-r^2_{01})^2}{N}.$$

If, however, there are no grounds for treating V_1 and V_2 as independent, except that r_{12} was insignificantly small in the data at hand, the proper set up is one in which a correlation between V_1 and V_2 is indicated as in Fig. 35.

FIG. 35

(83) $$r_{01} = p_{01} + p_{02}\, r_{12}$$

(84) $$r_{02} = p_{01}\, r_{12} + p_{02}$$

giving

(85) $$p_{01} = \frac{r_{01} - r_{02}\, r_{12}}{1 - r^2_{12}}.$$

Treating sampling errors as differentials

(86) $$\delta p_{01} = \frac{(1-r^2_{12})(\delta r_{01} - r_{02}\,\delta r_{12} - r_{12}\,\delta r_{02}) + 2(r_{01} - r_{02}\, r_{12})\, r_{12}\,\delta r_{12}}{(1-r^2_{12})^2}$$

In the present case, r_{12} (but not δr_{12}) is assumed to be zero in the sample at hand. Thus

(87) $$\delta p_{01} = \delta r_{01} - r_{02}\,\delta r_{12}$$

(88) $$\sigma^2_{p_{01}} = \sigma^2_{r_{01}} + r^2_{02}\,\sigma^2_{r_{12}} - 2 r_{02}\, m_{r_{01} r_{12}},$$

where $m_{r_{01} r_{02}}$ is the product moment of deviations of r_{01} and r_{12},

(89) $$m_{r_{01} r_{02}} = r_{02}(1-r^2_{01})(1-r^2_{12}) - \frac{r_{01} r_{12}}{2}(1 - r^2_{01} - r^2_{02} - r^2_{12} + 2 r_{01} r_{02} r_{12})$$

by the formula of Pearson and Filon. Again treating r_{12} as negligibly small,

(90) $$m_{r_{01} r_{02}} = r_{02}(1 - r^2_{01})$$

(91) $$\sigma^2_{p_{01}} = \frac{1}{N}\left[(1-r^2_{01})^2 + r^2_{02} - 2 r^2_{02}(1-r^2_{01})\right]$$

(92) $$= \sigma^2_{r_{01}} - \frac{r^2_{02}(1-2r^2_{01})}{N}.$$

This is smaller than the value of $\sigma^2_{p_{01}}$ obtained on the assumption of independence of V_1 and V_2 if r^2_{01} is less than ½, but larger for larger values of r^2_{01}.

If the correlation between the two known factors V_1 and V_2 of figure 35, is not negligible, the squaring of the full formula for δp_{01} and division by N, leads after some reduction to the formula

$$(93) \qquad \sigma^2_{p_{01}} = \frac{1}{N}\left[\frac{1-r^2_{0(12)}}{1-r^2_{12}} - p^2_{01}\left(1+r^2_{01}-2r^2_{0(12)}\right)\right].$$

A somewhat rough estimate of the standard errors in the analysis of birth weight of guinea pigs, page 179, can be made by this formula. The correlation between birth weight and size of litter was however based on larger numbers (3353) than the correlation involving gestation period (1317). Adopting the smaller numbers we find

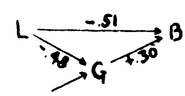


FIG. 36

$$p_{BL} = -.51 \pm .020$$

$$p_{BG} = +.30 \pm .022 .$$

While these estimates of the standard errors do not take cognizance of the approximation involved in substitution of gestation period for observed interval between litter (estimated $r_{GI} = .95$) they are sufficient to indicate that the calculated path coefficient can be relied upon as accurate to a first order, assuming the correctness of the set up.

The standard error of a path coefficient has not been worked out for systems in which one variable is represented as affected by more than two known variables. The standard error of the closely allied concrete regression coefficient is however well known and can be used in testing significance.

$$(94) \qquad \sigma^2_{c_{01}} = \frac{\sigma^2_{0.12\cdots m}}{\sigma^2_{1.2\cdots m}} = \frac{\sigma^2_0\left[1-r^2_{0(12\cdots m)}\right]}{\sigma^2_1\left[1-r^2_{1(2\cdots m)}\right]}$$

Since $p^2_{01} = c^2_{01}\dfrac{\sigma^2_1}{\sigma^2_0}$, the variance of the path coefficient can be written

(95) $\sigma^2_{P_{01}} = \dfrac{1-r^2_{0(12\cdots m)}}{1-r^2_{1(2\cdots m)}}$, if $\dfrac{\sigma_0^2}{\sigma_1^2}$ can be treated as constant. This probably gives fairly good approximation in any case and is so used by Brandt (1928). In the case of guinea pig weight discussed above, the correct formula gives a result a little smaller than this approximation.

It will be noted that the standard errors may take very high values if the independent variable under consideration (V_1) approaches complete determination by the others in the system, i.e. if $[1-r^2_{1(2\cdots m)}]$ approaches 0. In general, coefficients for paths leading from variables closely correlated with each other are subject to large standard errors. In making up a system, whether for prediction purposes or interpretation the aim should be to select factors closely correlated with the dependent variable but as nearly independent of each other as practicable.

If the dependent variable is completely determined by the specified factors $(r^2_{0(12\cdots m)}=1)$ the standard error of the concrete partial regression coefficient becomes zero. This is not the case with that of the path coefficient. Thus in the two factor case discussed above

(96) $\sigma^2_{P_{01}} = \dfrac{P_{01}^2(1-r_{01}^2)}{N} = \dfrac{(1-r_{01}^2)(1-r_{02}^2)}{N(1-r_{12}^2)}$ if $r^2_{0(12)}=1$.

More generally, if c_{01} can be treated as constant (as it can if $r^2_{0(12\cdots m)}=1$),

(97) $\sigma^2_{P_{01}} = c_{01}^2\,\sigma^2_{\left(\frac{\sigma_1}{\sigma_0}\right)} = c_{01}^2\,\dfrac{\sigma_1^2}{\sigma_0^2}\cdot\dfrac{(1-r_{01}^2)}{N} = \dfrac{P_{01}^2(1-r_{01}^2)}{N}$

which is in agreement with the preceding result.

Another simple set up, which is of interest is that in which three variables are arranged in chain sequence $(r_{02}=r_{01}\cdot r_{12})$. Here again the point of view makes a difference. If the above relation is merely an empirical one, the situation is mere-

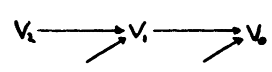


FIG. 37

ly a special case of that just discussed (the case in which $P_{02} = 0$, and $P_{01} = r_{01}$). By substitution we find

$$(98) \qquad \sigma^2_{P_{01}} = \frac{1}{N}\left[\frac{1-r^2_{01}}{1-r^2_{12}} - r^2_{01}\,(1-r^2_{01})\right]$$

$$(99) \qquad \sigma^2_{P_{12}} = \frac{1}{N}\left[1-r^2_{12}\right]^2$$

$$(100) \qquad \sigma^2_{P_{02}} = \frac{1}{N}\left[\frac{1-r^2_{01}}{1-r^2_{12}}\right].$$

If, however, V_1 is represented as the sole intermediary between V_0 and V_2 on theoretical grounds, the result is different. Two different determinations can be made of $\sigma^2_{P_{01}}$ and of $\sigma^2_{P_{12}}$, the reason being that more equations can be written than there are unknown path coefficients. From, $P_{01} = r_{01}$,

$$(101) \qquad \sigma^2_{P_{01}} = \frac{1}{N}\left(1-r^2_{01}\right)^2$$

. From $P_{01} = \frac{r_{02}}{r_{12}}$,

$$(102) \qquad \sigma^2_{P_{01}} = \frac{1}{N}\,\frac{(1-r^2_{01})(1-r^2_{02})}{r^2_{12}}$$

. From $P_{12} = r_{12}$,

$$(103) \qquad \sigma^2_{P_{12}} = \frac{1}{N}\left(1-r^2_{12}\right)^2$$

. From $P_{12} = \frac{r_{02}}{r_{01}}$,

$$(104) \qquad \sigma^2_{P_{12}} = \frac{1}{N}\,\frac{(1-r^2_{12})(1-r^2_{02})}{r^2_{01}}$$

Similarly two determinations can be made of $\sigma^2_{P_{012}}$. From $P_{012} = r_{02}$,

$$(105) \qquad \sigma^2_{P_{012}} = \frac{1}{N}\left[1-r^2_{02}\right]^2$$

. From $P_{012} = r_{01} \cdot r_{12}$,

$$(106) \qquad \sigma^2_{P_{012}} = \frac{1}{N}\left[(1-r^2_{02})^2 - (1-r^2_{01})(1-r^2_{12})\right].$$

With standard deviations calculated from two independent sets of data in each case, a combination estimate, smaller than either can be obtained from the formula

$$(107) \qquad \sigma^2_{Total} = \frac{1}{\Sigma \frac{1}{\sigma^2}}.$$

This illustrates the important principle that where there is a superfluity of equations for determining the path coefficients, the standard errors of these are correspondingly reduced. In the analysis of corn and hog correlations 42 path coefficients were found with which 510 correlations (and 4 cases of complete determination) were in agreement to the extent expected from their standard errors. Calculation of the standard errors of the path coefficients in this system seems out of the question, but it may safely be assumed that values of the order $\frac{1}{\sqrt{N}}$, which might be based on 42 equations are to be reduced by considerable amounts by the superfluity of data available.

There are some interesting contrasts in the standard errors given above. If r_{12}^2 is large, $\sigma_{P_{01}}^2$ may be large in the empirical system. But if the theory that X_1 is the only intermediary rests on adequate grounds, independent of the observed correlations, $\sigma_{P_{01}}^2$ may be small with large r_{12}^2 .

We will conclude with consideration of a set up like that used for the relation of supply and demand to price and quantity. It will be assumed first that the number of cases is large (a condition contrary to that found in the examples given). Differentiation of the 5 basic equations gives 5 equations expressing the relations between small deviations of the path coefficients and correlations.

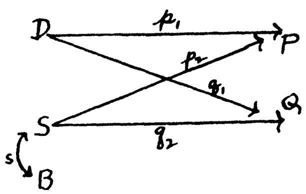


FIG. 38

(108) $p_1^2+p_2^2=1$ (113) $2p_1\,\delta p_1+2p_2\,\delta p_2=0$

(109) $q_1^2+q_2^2=1$ (114) $2q_1\,\delta q_1+2q_2\,\delta q_2=0$

(110) $p_1q_1+p_2q_2=r_{PQ}$ (115) $p_1\,\delta q_1+q_1\,\delta p_1+p_2\,\delta q_2+q_2\,\delta p_2=\delta r_{PQ}$

(111) $p_2 s=r_{BP}$ (116) $p_2\,\delta s+s\,\delta p_2=\delta r_{BP}$

(112) $q_2 s=r_{BQ}$ (117) $q_2\,\delta s+s\,\delta q_2=\delta r_{BQ}$.

Thus

(118) $$\delta p_1 = -\frac{p_2}{p_1}\,\delta p_2$$

(119) $$\delta q_1 = -\frac{q_2}{q_1}\,\delta q_2$$

(120) $$\delta q_2\left(p_2 - \frac{q_2}{q_1}p_1\right) + \delta p_2\left(q_2 - \frac{p_2}{p_1}q_1\right) = \delta r_{PQ}$$

(121) $$\delta q_2 \cdot p_2 - \delta p_2 \cdot q_2 = \frac{p_2}{s}\,\delta r_{BQ} - \frac{q_2}{s}\,\delta r_{BP}\,.$$

Solution of (120) and (121) as simultaneous equations gives expressions for δp_2 and δq_2 in terms of δr_{PQ}, δr_{BQ} & δr_{BP}, from which their squared standard errors can be found by taking the average squares. Letting A, B and C be the coefficients,

(122) $$\delta q_2 = A\,\delta r_{PQ} + B\,\delta r_{BQ} + C\,\delta r_{BP}$$

(123) $$\sigma_{q_2}^2 = A^2\sigma_{r_{PQ}}^2 + B^2\sigma_{r_{BQ}}^2 + C^2\sigma_{r_{BP}}^2 + 2AB\,m_{r_{PQ}r_{BQ}} + 2AC\,m_{r_{PQ}r_{BP}} + 2BC\,m_{r_{BQ}r_{BP}}$$

The product moments of the deviations of the correlation coefficients can be found by Pearson & Filon's formula cited on page 206.

The standard errors of δp_1 and δq_2 can be found at once with the help of equations (11) and (12) while that of δs can be found from (9) or (10) after expressing δp_2 (or δq_2) in terms of deviations of the known correlation coefficients.

The significance of the coefficients of elasticity is most easily investigated by taking these on scales in which the standard errors of the percentage deviation in price and quantity are taken as unity i.e. by finding the standard error of $\frac{q_1}{p_1}$ and of $\frac{q_2}{p_2}$ instead of $e = \frac{q_1\sigma_Q}{p_1\sigma_P}$ and $\eta = \frac{q_2\sigma_Q}{p_2\sigma_P}$ respectively. These standard errors can be found from the formula for the standard error of a ratio.

(124) $$\sigma_{\frac{q_1}{p_1}}^2 = \frac{q_1^2}{p_1^2}\left[\frac{\sigma_{q_1}^2}{q_1^2} + \frac{\sigma_{p_1}^2}{p_1^2} - 2\,\frac{m_{q_1p_1}}{q_1p_1}\right].$$

The product moments of the path coefficients can be obtained

by squaring equation (8) after expressing δp_1 and δq_1 in terms of δp_2 and δq_2 equation (13) or the converse.

The numbers of cases in the actual examples were not large enough to make the method a satisfactory one. The calculations have been carried through, however, with the results given below.

	Summer Pork 26 Years	Winter Pork 44 Years	Potatoes 19 Years
r_{PQ}	$-.63 \pm .12$	$-.68 \pm .08$	$-.85 \pm .06$
r_{PB}	$-.47 \pm .16$	$-.63 \pm .09$	$-.56 \pm .16$
r_{QB}	$-.64 \pm .12$	$+.83 \pm .05$	$+.65 \pm .14$
p_1	$+.69 \pm .20$	$+.66 \pm .11$	$+.50 \pm .26$
p_2	$-.73 \pm .19$	$-.76 \pm .09$	$-.86 \pm .15$
q_1	$+.13 \pm .24$	$+.11 \pm .10$	$+.02 \pm .27$
q_2	$+.99 \pm .03$	$+.99 \pm .01$	$+1.00 \pm .00$
s	$+.65 \pm .12$	$+.84 \pm .05$	$+.65 \pm .14$
q_1/p_1	$+.19 \pm .39$	$+.16 \pm .17$	$+.05 \pm .57$
q_2/p_2	$-1.36 \pm .38$	$-1.32 \pm .17$	$-1.16 \pm .21$
e	$+.13$	$+.11$	$+.03$
η	$-.94$	$-.88$	$-.82$

The most nearly satisfactory case is that of winter pork based on rather large primary correlations obtained from 44 years' experience, but even here, the standard error of q_1 is nearly as

large as q_1 itself. In the other cases, the standard error of q_1 is larger than q_1. The term δq_1^2 omitted in equation (7) is of the order of the term $2q_1 \delta q_1$ or larger, making this equation invalid. The other equations are not affected, at least to anything like as great an extent. An approximate solution can be obtained even though equation (7) is omitted, from the consideration that in this case in which q_1 is small, δq_2 must be very small and may be ignored. Thus

(125) $$\delta s = \delta r_{BQ}$$

(126) $$\delta p_2 = \frac{1}{s}\left[\delta r_{BP} - \frac{p_2}{q_2}\delta r_{BQ}\right]$$

(127) $$\delta q_1 = \frac{1}{p_1}\left[\delta r_{PQ} - q_2 \delta p_2\right]$$

(128) $$\delta p_1 = -\frac{p_2}{p_1}\delta p_2 .$$

The results are substantially the same as those obtained above, since the values assigned δq_2 were very small, even if not reliable. It may be safely concluded that winter pork at the large markets has very little elasticity of supply but a moderate elasticity of demand. The results for summer pork and for potatoes are in harmony with similar interpretations but are based on such inadequate numbers as to have little significance in themselves.

REFERENCES

Brandt, A. E., 1928—Calculation and use of the standard deviation of partial regression coefficients. Iowa St. College Jour. Sci. 2: 235–242.

Burks, B. S., 1928—The relative influence of nature and nurture upon mental development; a comparative study of foster-parent—foster-child resemblance and true parent—true child resemblance. 27th Yearbook of Nat. Soc. for Study of Education, 1928, Part I:219–316.

Calder, A., 1927—The role of inbreeding in the development of the Clydesdale breed of horse. Proc. Roy. Soc. Edinb. 47: 118–140.

Fisher, R. A., 1925—Statistical methods for Research Workers. 239 pp. Oliver & Boyd. Edinburgh.

Jennings, H. S., 1916—The numerical results of diverse systems of breeding. Genetics 1: 53–89.

Kelley, F. L., 1927—Statistical Method. 390 pp. The Macmillan Co. New York.

Krichewsky, S., 1927—Interpretation of Correlation Coefficients. Ministry of Public Works. Egypt. Physical Dept. Paper No. 22, Cairo.

Lush, J. L., 1930—The number of daughters necessary to prove a sire. Jour. Dairy Sci. 13: 209–220.

———1932—The amount and kind of inbreeding which has occurred in the development of breeds of livestock. Proc. 6th Internat. Congress of Genetics 2: 123–126.

McPhee, H. C., and S. Wright, 1925—Mendelian analysis of the pure breeds of live stock. III. The Shorthorns. Jour. Hered. 16: 205–215.

———1926—Mendelian analysis of the pure breeds of live stock. IV. The British Dairy Shorthorns. Jour. Hered. 17: 397–401.

Minot, C. S., 1891—Senescence and rejuvenation. Jour. Physiol. 12: 97–153.

Niles, H. E., 1922—Correlation, Causation and Wright's theory of path coefficients. Genetics 7: 258–273.

———1923—The method of path coefficients, an answer to Wright. Genetics 8: 256–260.

Smith, A. D. B., 1926—Inbreeding in cattle and horses. Eugen. Rev. 14: 189–204.

Wright, P. G., 1928—The tariff on animal and vegetable oils. 347 pp. The Macmillan Co., New York.

Wright, S., 1918—On the nature of size factors. Genetics 3: 367–374.

———1920—The relative importance of heredity and environment in determining the piebald pattern of guinea pigs. Proc. Nat. Acad. Sci. 6: 320–332.

———1921a—Correlation and Causation. Jour. Ag. Res. 20: 557–585.

———1921b—Systems of mating. Genetics 6: 111–178.

———1922 —Coefficients of inbreeding and relationship. Am. Nat. 56: 330–338.

———1923a—The theory of path coefficients—a reply to Niles' criticism. Genetics 8: 239–255.

———1923b—Mendelian analysis of the pure breeds of livestock. I. The measurement of inbreeding and relationship. Jour. Hered. 14: 339–348. II. The Duchess family of Shorthorns as bred by Thomas Bates. Jour. Hered. 14: 405–422.

———1925a—Corn and hog correlations. Bull. No. 1300, 60 pp. U. S. Dept. of Agric.

———1926a—A frequency curve adapted to variation in percentage occurrence. Jour. Amer. Stat. Assoc. 21: 162–178.

———1926b—Effects of age ot parents on characteristics of the guinea pig. Amer. Nat. 60: 552–559.

———1927 —The effects in combination of the major color-factors of the guinea pig. Genetics 12: 530–569.

———1931a—Statistical methods in biology. Jour. Amer. Stat. Ass. Supplement. Papers and Proceedings of the 92nd annual meeting. 26: 155–163.

———1931b—Evolution in mendelian populations. Genetics 16: 97–159.

———1932a—On the evaluation of dairy sires. Proc. Amer. Soc. Animal Prod. 1932: 71–78.

———1932b—General, group and special size factors. Genetics 17: 603–619.

Wright, S., and H. C. McPhee, 1925—An approximate method of calculating coefficients of inbreeding and relationship from livestock pedigrees. Jour. Ag. Res. 31: 377–383.

Wright, Sewall, 1933a—Inbreeding and homozygosis. Proc. Nat. Acad. Sci. 19: 411–420.

———1933b—Inbreeding and recombination. Proc. Nat. Acad. Sci. 19:420–433.

Sewall Wright